GenCore version 5.1.8

OM nucleic - nucleic search, using sw model

Run on: May 6, 2006, 19:28:56 ; Search time 2414 Seconds
(without alignments)
8453.521 Million cell updates/sec

Title: US-10-009-579A-5_COPY_3188_3546
Perfect score: 359
Sequence: 1 cccggctaattttgtatctt.....ttttttatagtgttctggaa 359

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	359	100.0	4069	8	AY148099	AY148099 Homo sapi
	2	359	100.0	4282	6	AX254778	AX254778 Sequence
	3	359	100.0	171987	8	AC079775	AC079775 Homo sapi
	4	132.8	37.0	56173	8	AC011390	AC011390 Homo sapi
c	5	132.8	37.0	130985	8	HSA243213	AJ243213 Homo sapi
c	6	132.8	37.0	153553	8	AC008627	AC008627 Homo sapi
c	7	132.4	36.9	84710	8	HSBA358N2	AL121583 Human DNA
c	8	130.8	36.4	176075	8	AC017079	AC017079 Homo sapi
c	9	127.4	35.5	162509	8	AL137852	AL137852 Human DNA
	10	127	35.4	568	6	BD152002	BD152002 Primer fo
	11	127	35.4	568	6	AX871940	AX871940 Sequence
c	12	126.8	35.3	139255	8	AL512642	AL512642 Human DNA
	13	126.8	35.3	148285	14	AL451053	AL451053 Homo sapi
c	14	126.8	35.3	192096	8	AL590133	AL590133 Human DNA
c	15	126	35.1	110459	8	AC117378	AC117378 Homo sapi
c	16	126	35.1	113802	8	AL137003	AL137003 Human DNA
	17	126	35.1	132070	8	AC003663	AC003663 Homo sapi
	18	126	35.1	160298	14	AC067818	AC067818 Homo sapi
	19	126	35.1	161139	8	AC132812	AC132812 Homo sapi
	20	125.8	35.0	138604	8	AC104826	AC104826 Homo sapi
c	21	125.8	35.0	160457	8	AC006016	AC006016 Homo sapi
	22	125.8	35.0	169334	14	AC025406	AC025406 Homo sapi
	23	125.8	35.0	172464	14	AC140889	AC140889 Homo sapi
	24	125.8	35.0	201312	14	AC067900	AC067900 Homo sapi
	25	125.6	35.0	54666	8	AC073487	AC073487 Homo sapi
c	26	125.6	35.0	105736	14	AC090678	AC090678 Homo sapi
	27	125.6	35.0	197137	14	AC024384	AC024384 Homo sapi
	28	125.4	34.9	88119	8	AL450325	AL450325 Human DNA
c	29	125.4	34.9	175067	14	AC040898	AC040898 Homo sapi
c	30	125.4	34.9	180754	14	AC148963	AC148963 Papio anu
	31	125	34.8	61371	14	AC105135	AC105135 Homo sapi
	32	125	34.8	190579	14	AC018736	AC018736 Homo sapi
	33	125	34.8	190836	8	AC105910	AC105910 Homo sapi
c	34	124.8	34.8	52468	14	AC022628	AC022628 Homo sapi
c	35	124.8	34.8	186563	8	AC104452	AC104452 Homo sapi
c	36	124.8	34.8	186925	14	AC087503	AC087503 Homo sapi
c	37	124.6	34.7	161780	8	AC145893	AC145893 Pan trogl
c	38	124.4	34.7	68470	14	AC116011	AC116011 Homo sapi
c	39	124.4	34.7	85123	8	AC125616	AC125616 Homo sapi
c	40	124.4	34.7	137242	8	AC063943	AC063943 Homo sapi
c	41	124.4	34.7	183228	8	AC010547	AC010547 Homo sapi
	42	124.4	34.7	203050	14	HS44N10	Z97197 Homo sapien
	43	124.4	34.7	206943	14	AC138848	AC138848 Homo sapi
	44	124.4	34.7	208497	14	AC090584	AC090584 Homo sapi
c	45	124.4	34.7	209512	8	CNS018OV	AL109758 Human chr

ALIGNMENTS

RESULT 1
AY148099
LOCUS AY148099 4069 bp DNA linear PRI 12-NOV-2002
DEFINITION Homo sapiens EGP2 (TACSTD1) gene, promoter region and 5'UTR.
ACCESSION AY148099
VERSION AY148099.1 GI:24935271
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4069)
AUTHORS McLaughlin,P.M.J., Trzpis,M., Kroesen,B.-J., Helfrich,W., Terpstra,P., Ruiters,M.H.C., de Leij,L.F.M.H. and Harmsen,M.C.
TITLE Use of the EGP-2 promoter for targeted expression of heterologous genes in carcinoma-derived cell lines
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4069)
AUTHORS McLaughlin,P.M.J., Trzpis,M., Kroesen,B.-J., Helfrich,W., Terpstra,P., Ruiters,M.H.C., de Leij,L.F.M.H. and Harmsen,M.C.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-2002) Path-LabMed-Mol-Biol, University Groningen, Hanzeplein 1, Groningen 9713GZ, The Netherlands
FEATURES Location/Qualifiers
source 1. .4069
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2p16.3"
gene 1. .>4069
/gene="TACSTD1"
/note="synonyms: EGP2, GA733-2"
promoter 1. .3974
/gene="TACSTD1"
mRNA 3975. .>4069
/gene="TACSTD1"
/product="EGP2"
/note="putative transcription start"
5'UTR 3975. .>4069

```
                     /gene="TACSTD1"
ORIGIN

  Query Match             100.0%;  Score 359;  DB 8;  Length 4069;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-69;
  Matches  359;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3185 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 3244

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3245 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 3304

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3305 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 3364

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3365 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 3424

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTTTTCTTTCTAATTTCAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3425 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTTTTCTTTCTAATTTCAA 3484

Qy        301 AGGAGTATAATTAAATTGCCAGGTAAAAGCTCAAAGGTCTTTTTTATAGTGTTCTGGAA 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3485 AGGAGTATAATTAAATTGCCAGGTAAAAGCTCAAAGGTCTTTTTTATAGTGTTCTGGAA 3543


RESULT 2
AX254778
LOCUS       AX254778               4282 bp    DNA     linear   PAT 10-OCT-2001
DEFINITION  Sequence 5 from Patent WO0171015.
ACCESSION   AX254778
VERSION     AX254778.1  GI:16074440
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   de Leij,L.F., Ruiters,M.H., Mclaughlin,P.M., Harmsen,M.C., van der
            Molen,H., Terpstra,P. and Dokter,W.H.
  TITLE     Non-squamous epithelium-specific transcription
  JOURNAL   Patent: WO 0171015-A 5 27-SEP-2001;
            Rijksuniversiteit Groningen (NL)
FEATURES             Location/Qualifiers
     source          1. .4282
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    1. .4282
                     /note="'EGP-2 promotor sequence from -3967 to +315'"
ORIGIN

  Query Match             100.0%;  Score 359;  DB 6;  Length 4282;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-69;
  Matches  359;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3188 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 3247

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3248 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 3307

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3308 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 3367

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3368 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 3427

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTTTTCTTTCTAATTTCAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3428 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTTTTCTTTCTAATTTCAA 3487

Qy        301 AGGAGTATAATTAAATTGCCAGGTAAAAGCTCAAAGGTCTTTTTTATAGTGTTCTGGAA 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3488 AGGAGTATAATTAAATTGCCAGGTAAAAGCTCAAAGGTCTTTTTTATAGTGTTCTGGAA 3546


RESULT 3
AC079775
LOCUS       AC079775               171987 bp    DNA     linear   PRI 21-APR-2005
DEFINITION  Homo sapiens BAC clone RP11-295P2 from 2, complete sequence.
ACCESSION   AC079775
VERSION     AC079775.6  GI:19848453
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 171987)
  AUTHORS   Elliott,G., Doebber,A., Belter,E. and Haakenson,W.
  TITLE     The sequence of Homo sapiens BAC clone RP11-295P2
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 171987)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-SEP-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 171987)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 171987)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAY-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 171987)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Mar 30, 2002 this sequence version replaced gi:16924154.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0295P02
            --------------.

            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.
```

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-261E8; the clone sequenced to the right is RP11-436K12. Actual start of this clone is at base position 1 of RP11-295P2; actual end is at base position 171987 of RP11-295P2.

FEATURES Location/Qualifiers

source 1. .171987
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-295P2"
/clone_lib="RPCI-11"

misc_feature 64058. .65106
/note="CpG_island (%GC=74.0, o/e=0.87, #CpGs=126)"

gene 64313. .82007
/gene="TACSTD1"

mRNA join(64313. .64570,68452. .68559,68797. .69037,70223. .70288, 72003. .72066,73942. .74043,74758. .74958,80155. .80199, 81561. .82007)
/gene="TACSTD1"

CDS join(64495. .64570,68452. .68559,68797. .69037,70223. .70288, 72003. .72066,73942. .74043,74758. .74958,80155. .80199, 81561. .81602)
/gene="TACSTD1"
/note="Homo sapiens tumor-associated calcium signal transducer 1 (TACSTD1), mRNA.; H_NH0295P02.1
This gene was based on gi(4505058)"
/codon_start=1
/product="unknown"
/protein_id="AAY15095.1"
/db_xref="GI:62822547"
/translation="MAPPQVLAFGLLLAAATATFAAAQEECVCENYKLAVNCFVNNNR QCQCTSVGAQNTVICSKLAAKCLVMKAEMNGSKLGRRAKPEGALQNNDGLYDPDCDES GLFKAKQCNGTSMCWCVNTAGVRRTDKDTEITCSERVRTYWIIIELKHKAREKPYDSK SLRTALQKEITTRYQLDPKFITSILYENNVITIDLVQNSSQKTQNDVDIADVAYYFEK DVKGESLFHSKKMDLTVNGEQLDLDPGQTLIYYVDEKAPEFSMQGLKAGVIAVIVVVV IAVVAGIVVLVISRKKRMAKYEKAEIKEMGEMHRELNA"

misc_feature 65321. .65560
/gene="TACSTD1"
/note="CpG_island (%GC=58.8, o/e=0.81, #CpGs=17)"

misc_feature 97918. .98641
/note="CpG_island (%GC=67.1, o/e=0.92, #CpGs=76)"

gene 98113. .166051
/gene="MSH2"

mRNA join(98113. .98391,103390. .103544,105083. .105361, 107403. .107549,109258. .109407,111285. .111418, 124731. .124930,140537. .140646,158020. .158143, 161647. .161797,165954. .166051)
/gene="MSH2"

CDS join(98181. .98391,103390. .103544,105083. .105361, 107403. .107549,109258. .109407,111285. .111418, 124731. .124930,140537. .140646,158020. .158143, 161647. .161797,165954. .>166051)
/gene="MSH2"
/note="Homo sapiens mutS homolog 2, colon cancer, nonpolyposis type 1 (E. coli) (MSH2), mRNA.; H_NH0295P02.2
This gene was based on gi(4557760)
Continues as H_NH0436K12.1"
/codon_start=1
/product="unknown"
/protein_id="AAY15096.1"
/db_xref="GI:62822548"
/translation="MAVQPKETLQLESAAEVGFVRFFQGMPEKPTTTVRLFDRGDFYT AHGEDALLAAREVFKTQGVIKYMGPAGAKNLQSVVLSKMNFESFVKDLLLVRQYRVEV YKNRAGNKASKENDWYLAYKASPGNLSQFEDILFGNNDMSASIGVVGVKMSAVDGQRQ VGVGYVDSIQRKLGLCEFPDNDQFSNLEALLIQIGPKECVLPGGETAGDMGKLRQIIQ RGGILITERKKADFSTKDIYQDLNRLLKGKKGEQMNSAVLPEMENQVAVSSLSAVIKF LELLSDDSNFGQFELTTFDFSQYMKLDIAAVRALNLFQGSVEDTTGSQSLAALLNKCK TPQGQRLVNQWIKQPLMDKNRIEERLNLVEAFVEDAELRQTLQEDLLRRFPDLNRLAK KFQRQAANLQDCYRLYQGINQLPNVIQALEKHEGKHQKLLLAVFVTPLTDLRSDFSKF QEMIETTLDMDQVENHEFLVKPSFDPNLSELREIMNDLEKKMQSTLISAARDLGLDPG KQIKLDSSAQFGYYFRVTCKEEKVLRNNKNFSTVDIQKNGVKFTNSKLTSLNEEYTKN KTEYEEAQDAIVKEIVNISS"

ORIGIN

Query Match 100.0%; Score 359; DB 8; Length 171987;
Best Local Similarity 100.0%; Pred. No. 8.4e-69;
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63527 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 63586

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63587 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 63646

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63647 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 63706

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63707 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 63766

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTTTTCTTTCTAATTTCAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63767 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTTTTCTTTCTAATTTCAA 63826

Qy        301 AGGAGTATAATTAAATTGCCAGGTAAAAGCTCAAAGGTCTTTTTTATAGTGTTCTGGAA 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63827 AGGAGTATAATTAAATTGCCAGGTAAAAGCTCAAAGGTCTTTTTTATAGTGTTCTGGAA 63885
```

RESULT 4
AC011390
LOCUS AC011390 56173 bp DNA linear PRI 26-JAN-2002
DEFINITION Homo sapiens chromosome 5 clone CTB-176I15, complete sequence.
ACCESSION AC011390
VERSION AC011390.3 GI:18376861
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 56173)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 56173)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 56173)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission

```
  JOURNAL   Submitted (26-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jan 26, 2002 this sequence version replaced gi:7710540.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES             Location/Qualifiers
     source          1. .56173
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTB-176I15"
ORIGIN

  Query Match             37.0%;  Score 132.8;  DB 8;  Length 56173;
  Best Local Similarity   76.8%;  Pred. No. 4.2e-19;
  Matches  175;  Conservative    0;  Mismatches   52;  Indels    1;  Gaps    1;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              || ||||||||||||||| |||||||||||||||| ||| ||||||||||||||||||||
Db      43992 CCTGGCTAATTTTGTAT-TTTTAGTAGAGACGGGGTTTCTCCATGTTGGTCAGGCTGGTC 44050

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              |||||||||  ||||||||||||||||||||||||||||||||||||||| |||||||||
Db      44051 TCGAACTTCTGACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 44110

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              || |||||||||||| | ||| || |  |  ||||  |||||  |           |
Db      44111 GCATGAGCCACCGCGCCCGGCCAGGAATGAAGTTTTGATACATGCTACCACATGGATGAA 44170

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACA 228
               ||||  || |    |||  ||| | |||| | ||| |   | | |||
Db      44171 CCTTAAAAATAAATGAAACAAATAAATAAGTGAAATAAATCACACACA 44218


RESULT 5
HSA243213/c
LOCUS       HSA243213             130985 bp    DNA     linear   PRI 02-FEB-2000
DEFINITION  Homo sapiens partial 5-HT4 receptor gene, exons 2 to 5.
ACCESSION   AJ243213
VERSION     AJ243213.1  GI:6900061
KEYWORDS    5-HT4 gene; 5-HT4 receptor.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Bender,E., Pindon,A., van Oers,I., Zhang,Y.B., Gommeren,W.,
            Verhasselt,P., Jurzak,M., Leysen,J. and Luyten,W.
  TITLE     Structure of the human serotonin 5-HT4 receptor gene and cloning of
            a novel 5-HT4 splice variant
  JOURNAL   J. Neurochem. 74 (2), 478-489 (2000)
   PUBMED   10646498
REFERENCE   2  (bases 1 to 130985)
  AUTHORS   Bender,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-1999) Bender E., Functional Genomics, Janssen
            Pharmaceutica, Turnhoutseweg 30, B-2340 Beerse, BELGIUM
FEATURES             Location/Qualifiers
     source          1. .130985
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="pBeloBAC11-228K23"
                     /clone_lib="Human genomic DNA in pBeloBAC11 (Research
                     Genetics, Huntsville, AL, U.S.A.)"
     gene            9622. .94002
                     /gene="5-HT4"
     mRNA            join(<9622. .9747,11016. .11216,36537. .36690,38920. .38961,
                     49859. .50427,76574. .76664,78238. .78397,83122. .83137,
                     93942. .94002)
                     /gene="5-HT4"
     CDS             join(<9622. .9747,11016. .11216,36537. .36690,38920. .38961,
                     49859. .50427,76574. .76664)
                     /gene="5-HT4"
                     /function="G protein coupled receptor"
                     /codon_start=2
                     /product="5-HT4 receptor"
                     /protein_id="CAB71316.1"
                     /db_xref="GI:6900062"
                     /db_xref="GOA:Q13639"
                     /db_xref="UniProt/Swiss-Prot:Q13639"
                     /translation="SEEGFGSVEKVVLLTFLSTVILMAILGNLLVMVAVCWDRQLRKI
                     KTNYFIVSLAFADLLVSVLVMPFGAIELVQDIWIYGEVFCLVRTSLDVLLTTASIFHL
                     CCISLDRYYAICCQPLVYRNKMTPLRIALMLGGCWVIPTFISFLPIMQGWNNIGIIDL
                     ERSLNQGLGQDFHAIEKRKFNQNSNSTYCVFMVNKPYAITCSVVAFYIPFLLMVLAYY
                     RIYVTAKEHAHQIQMLQRAGASSESRPQSADQHSTHRMRTETKAAKTLCIIMGCFCLC
                     WAPFFVTNIVDPFIDYTVPGQVWTAFLWLGYINSGLNPFLYAFLNKSFRRAFLIILCC
                     DDERYRRPSILGQTVPCSTTTINGSTHVLRDAVECGGQWESQCHPPATSPLVAAQPSD
                     T"
     exon            9622. .9747
                     /gene="5-HT4"
                     /number=2
     intron          9748. .11015
                     /gene="5-HT4"
                     /number=2
     exon            11016. .11216
                     /gene="5-HT4"
                     /number=3
     intron          11217. .36536
                     /gene="5-HT4"
                     /number=3
     exon            36537. .36690
                     /gene="5-HT4"
                     /number=4
     intron          36691. .38919
                     /gene="5-HT4"
                     /number=4
     exon            38920. .38961
                     /gene="5-HT4"
                     /note="number h"
     intron          38962. .49858
                     /gene="5-HT4"
                     /note="number h"
     exon            49859. .50427
                     /gene="5-HT4"
                     /number=5
     intron          50428. .76573
                     /gene="5-HT4"
                     /number=5
     exon            76574. .76664
                     /gene="5-HT4"
                     /note="number b"
     intron          76665. .78237
                     /gene="5-HT4"
                     /note="number b"
     exon            78238. .78397
                     /gene="5-HT4"
                     /note="number c"
     intron          78398. .83121
                     /gene="5-HT4"
                     /note="number c"
     exon            83122. .83137
                     /gene="5-HT4"
                     /note="number d"
     intron          83138. .93941
                     /gene="5-HT4"
                     /note="number d"
     exon            93942. .94002
                     /gene="5-HT4"
                     /note="number gef"
```

```
ORIGIN

  Query Match             37.0%;  Score 132.8;  DB 8;  Length 130985;
  Best Local Similarity   76.8%;  Pred. No. 4e-19;
  Matches  175;  Conservative    0;  Mismatches   52;  Indels    1;  Gaps    1;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              || ||||||||||||||| ||||||||||||||| ||| ||||||||||||||||||||||
Db       5650 CCTGGCTAATTTTGTAT-TTTTAGTAGAGACGGGGTTTCTCCATGTTGGTCAGGCTGGTC 5592

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              |||||||||  |||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       5591 TCGAACTTCTGACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 5532

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              || |||||||||||||| | ||| || |  |  |||| |||||  |             |
Db       5531 GCATGAGCCACCGCGCCCGGCCAGGAATGAAGTTTTGATACATGCTACCACATGGATGAA 5472

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACA 228
              ||||  || |    |||  ||| | |||| | ||| |   | | |||
Db       5471 CCTTAAAAATAAATGAAACAAATAAATAAGTGAAATAAATCACACACA 5424


RESULT 6
AC008627/c
LOCUS       AC008627              153553 bp    DNA     linear   PRI 11-APR-2002
DEFINITION  Homo sapiens chromosome 5 clone CTB-160O22, complete sequence.
ACCESSION   AC008627
VERSION     AC008627.7  GI:20136896
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 153553)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 153553)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 153553)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-APR-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 11, 2002 this sequence version replaced gi:12830083.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
            NOTE:  This insert is not the entire sequence of the clone (entire
            sequence is 197.3kb). It is clipped at the overlap with AC011390.
            The number of bases overlapped is 138359.
FEATURES             Location/Qualifiers
     source          1. .153553
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTB-160O22"
ORIGIN

  Query Match             37.0%;  Score 132.8;  DB 8;  Length 153553;
  Best Local Similarity   76.8%;  Pred. No. 4e-19;
  Matches  175;  Conservative    0;  Mismatches   52;  Indels    1;  Gaps    1;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              || ||||||||||||||| ||||||||||||||| ||| ||||||||||||||||||||||
Db     150536 CCTGGCTAATTTTGTAT-TTTTAGTAGAGACGGGGTTTCTCCATGTTGGTCAGGCTGGTC 150478

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              |||||||||  |||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     150477 TCGAACTTCTGACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 150418

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              || |||||||||||||| | ||| || |  |  |||| |||||  |             |
Db     150417 GCATGAGCCACCGCGCCCGGCCAGGAATGAAGTTTTGATACATGCTACCACATGGATGAA 150358

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACA 228
              ||||  || |    |||  ||| | |||| | ||| |   | | |||
Db     150357 CCTTAAAAATAAATGAAACAAATAAATAAGTGAAATAAATCACACACA 150310


RESULT 7
HSBA358N2/c
LOCUS       HSBA358N2             84710 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-358N2 on chromosome 20 Contains
            the 5' end of the ASXL1 gene for additional sex combs like 1
            (Drosophila) and a CpG island, complete sequence.
ACCESSION   AL121583
VERSION     AL121583.25  GI:9588417
KEYWORDS    HTG; ASXL1; CpG island; KIAA0978.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 84710)
  AUTHORS   Lloyd,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 31, 2000 this sequence version replaced gi:9230894.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 20, constructed by the Sanger Centre Chromosome 20
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr20
            RP11-358N2 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .84710
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="RP11-358N2"
                     /clone_lib="RPCI-11.2"
     misc_feature    1
                     /note="Clone_left_end: RP11-358N2"
```

```
gene            join(7902. .8216,15768. .15850,17111. .17113,18399. .18507,
                77512. .77632,77709. .77806,78722. .78815,79285. .79437,
                80705. .80868,80967. .81063,82264. .82369,82668. .83301,
                83816. .84710,AL034550.31:101. .4093)
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-001"
mRNA            join(7902. .8216,15768. .15850,17111. .17113,18399. .18507,
                77512. .77632,77709. .77806,78722. .78815,79285. .79437,
                80705. .80868,80967. .81063,82264. .82369,82668. .83301,
                83816. .84710,AL034550.31:101. .4093)
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-001"
                /note="match: ESTs: AI450494.1 AI632803.1 AL705053.1
                AW500268.1 AW503625.1 AW957415.1 BE859912.1 BE867570.1
                BE952301.1 BF848475.1 BI094451.1 BM456370.1 BM676077.1
                BM985410.1 BM988291.1 BM991942.1 BQ006322.1 BQ064719.1
                BU501765.1 BU633817.1 BU675124.1 BU685659.1 BU686340.1
                CA313728.1 CA417500.1 CA417510.1
                match: cDNAs: AB023195.2 AJ438952.1 AK025756.1 AL117518.1
                AL117647.1"
CDS             join(8160. .8216,15768. .15850,17111. .17113,18399. .18507,
                77512. .77632,77709. .77806,78722. .78815,79285. .79437,
                80705. .80868,80967. .81063,82264. .82369,82668. .83301,
                83816. .84710,AL034550.31:101. .2112)
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-001"
                /standard_name="OTTHUMP00000030592"
                /note="match: proteins: Q9UFJ0 Q9UFP8 Q9Y2I4"
                /codon_start=1
                /protein_id="CAI40414.1"
                /db_xref="GI:57162101"
                /translation="MKDKQKKKKERTWAEAARLVLENYSDAPMTPKQILQVIEAEGLK
                EMRSGTSPLACLNAMLHSNSRGGEGLFYKLPGRISLFTLKKDALQWSRHPATVEGEEP
                EDTADVESCGSNEASTVSGENDVSLDETSSNASCSTESQSRPLSNPRDSYRASSQANK
                QKKKTGVMLPRVVLTPLKVNGAHVESASGFSGCHADGESGSPSSSSSGSLALGSAAIR
                GQAEVTQDPAPLLRGFRKPATGQMKRNRGEEIDFETPGSILVNTNLRALINSRTFHAL
                PSHFQQQLLFLLPEVDRQVGTDGLLRLSSSALNNEFFTHAAQSWRERLADGEFTHEMQ
                VRIRQEMEKEKKVEQWKEKFFEDYYGQKLGLTKEESLQQNVGQEEAEIKSGLCVPGES
                VRIQRGPATRQRDGHFKKRSRPDLRTRARRNLYKKQESEQAGVAKDAKSVASDVPLYK
                DGEAKTDPAGLSSPHLPGTSSAAPDLEGPEFPVESVASRIQAEPDNLARASASPDRIP
                SLPQETVDQEPKDQKRKSFEQAASASFPEKKPRLEDRQSFRNTIESVHTEKPQPTKEE
                PKVPPIRIQLSRIKPPWVVKGQPTYQICPRIIPTTESSCRGWTGARTLADIKARALQV
                RGARGHHCHREAATTAIGGGGGPGGGGGGATDEGGGRGSSSGDGGEACGHPEPRGGPS
                TPGKCTSDLQRTQLLPPYPLNGEHTQAGTAMSRARREDLPSLRKEESCLLQRATVGLT
                DGLGDASQLPVAPTGDQPCQALPLLSSQTSVAERLVEQPQLHPDVRTECESGTTSWES
                DDEEQGPTVPADNGPILSLVGDDTLEKGTGQALDSHPTMKDPVNVTPSSTPESSPTDC
                LQNRAFDDELGLGGSCPPMRESDTRQENLKTKALVSNSSLHWIPIPSNDEVVKQPKPE
                SREHIPSVEPQVGEEWEKAAPTPPALPGDLTAEEGLDPLDSLTSLWTVPSRGGSDSNG
                SYCQQVDIEKLKINGDSEALSPHGESTDTASDFEGHLTEDSSEADTREAAVTKGSSVD
                KDEKPNWNQSAPLSKVNGDMRLVTRTDGMVAPQSWVSRVCAVRQKIPDSLLLASTEYQ
                PRAVCLSMPGSSVEATNPLVMQLLQGSLPLEKVLPPAHDDSMSESPQVPLTKDQSHGS
                LRMGSLHGLGKNSGMVDGSSPSSLRALKEPLLPDSCETGTGLARIEATQAPGAPQKNC
                KAVPSFDSLHPVTNPITSSRKLEEMDSKEQFSSFSCEDQKEVRAMSQDSNSNAAPGKS
                PGDLTTSRTPRFSSPNVISFGPEQTGRALGDQSNVTGQGKKLFGSGNVAATLQRPRPA
                DPMPLPAEIPPVFPSGKLGPSTNSMSGGVQTPREDWAPKPHAFVGSVKNEKTFVGGPL
                KANAENRKATGHSPLELVGHLEGMPFVMDLPFWKLPREPGKGLSEPLEPSSLPSQLSI
                KQAFYGKLSKLQLSSTSFNYSSSSPTFPKGLAGSVVQLSHKANFGASHSASLSLQMFT
                DSSTVESISLQCACSLKAMIMCQGCGAFCHDDCIGPSKLCVLCLVVR"
gene            join(8195. .8216,15768. .15850,17111. .17113,18399. .18507,
                21548. .21614)
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-003"
mRNA            join(8195. .8216,15768. .15850,17111. .17113,18399. .18507,
                21548. .21614)
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-003"
                /note="match: ESTs: BF000474.1"
gene            join(8949. .9175,15768. .15850,17111. .17113,18399. .18507,
                21548. .21937)
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-002"
mRNA            join(8949. .9175,15768. .15850,17111. .17113,18399. .18507,
                21548. .21937)
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-002"
                /note="match: ESTs: AI034028.1 AI829840.1 BM929970.1
                BU682051.1"
CDS             join(9131. .9175,15768. .15850,17111. .17113,18399. .18507,
                21548. .21553)
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-002"
                /standard_name="OTTHUMP00000030593"
                /codon_start=1
                /protein_id="CAI40415.1"
                /db_xref="GI:57162102"
                /db_xref="UniProt/TrEMBL:Q5JWS8"
                /translation="MGVVNFVFRASEDLQVLENYSDAPMTPKQILQVIEAEGLKEMRS
                GTSPLACLNAMLHSNSRGGEGLFYKLPGRISLFTLKV"
gene            join(15768. .15850,17111. .17113,18399. .18507,77512. .77806,
                78722. .79304)
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-004"
mRNA            join(15768. .15850,17111. .17113,18399. .18507,77512. .77806,
                78722. .79304)
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-004"
                /note="match: ESTs: BM906623.1"
polyA_signal    21913. .21918
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-002"
polyA_site      21937
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-002"
misc_feature    56954
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-004"
                /note="Clone_right_end: RP1-316I5"
misc_feature    84611
                /gene="ASXL1"
                /locus_tag="RP11-358N2.1-001"
                /note="Clone_left_end: RP5-1184F4"
ORIGIN

  Query Match           36.9%;  Score 132.4;  DB 8;  Length 84710;
  Best Local Similarity  65.6%;  Pred. No. 5e-19;
  Matches  193;  Conservative    0;  Mismatches  101;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              || |||||||||||      |||||||||||||||| ||| | ||||||||| ||||||||||
Db      11309 CCTGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTC 11250

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              |||||||||   |||||||||||||| | |||||||||||||||||| ||| |||||||||
Db      11249 TCGAACTTCTGGCCTCAGGTGATCCGTCTGCCTCGGCCTCCCAAAGAGCTGGGATTACAG 11190

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||||| || | ||||||||          | |  | |||  | | |   |
Db      11189 GCGTGAGCCACCGTGCCCGGCCTGGGACTCTTAAAAAGTTCCCCCTCACCTACCATCCAG 11130

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
              |  ||  |||  ||   | |||| | ||     ||      |   || |  |      | |
Db      11129 TCCCATTCAAAGAAATGGAAAATTCTAAAACAAAAAACACACACAAGATACTCAAATTGC 11070

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTTTTCTTTCTAA 294
              | | |  |   |  | ||||     || ||||||||||||| ||| | |
Db      11069 ATATCACATTCTCAACAAGTCCTTTTTCTTTTTTTCTTTTTTTTTTTTTTTTGA 11016


RESULT 8
AC017079/c
LOCUS       AC017079              176075 bp    DNA     linear   PRI 21-APR-2005
DEFINITION  Homo sapiens BAC clone RP11-462M9 from 2, complete sequence.
ACCESSION   AC017079
VERSION     AC017079.5  GI:13625506
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 176075)
  AUTHORS   Hawkins,M., Maupin,R., Le,T. and Belter,E.
  TITLE     The sequence of Homo sapiens BAC clone RP11-462M9
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 176075)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-DEC-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 176075)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 176075)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-AUG-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 176075)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-NOV-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 176075)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Apr 15, 2001 this sequence version replaced gi:13431206.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0462M09
            --------------.

            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:

            The clone sequenced to the right is RP11-480N9.  Actual start of
            this clone is at base position 1 of RP11-462M9; actual end is at
            base position 176075 of RP11-462M9.

            The sequence H_NH0462M09 from base positions 164804 to 164876 and
            165220 to 165298 are represented by sequence derived from PCR  on
            BAC DNA.

FEATURES             Location/Qualifiers
     source          1. .176075
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-462M9"
                     /clone_lib="RPCI-11"
     gene            2003. .37935
                     /gene="UGCGL1"
     mRNA            join(2003. .2101,2906. .2929,3766. .3860,6028. .6234,
                     6974. .7105,7652. .7782,11244. .11339,16802. .16904,
                     17719. .17798,19087. .19276,20295. .20453,21287. .21372,
                     23328. .23412,24346. .24434,24982. .25083,26300. .26384,
                     27404. .27586,28644. .28804,30189. .30303,33197. .33313,
                     33963. .34128,36231. .37935)
                     /gene="UGCGL1"
     CDS             join(<2003. .2101,2906. .2929,3766. .3860,6028. .6234,
                     6974. .7105,7652. .7782,11244. .11339,16802. .16904,
                     17719. .17798,19087. .19276,20295. .20453,21287. .21372,
                     23328. .23412,24346. .24434,24982. .25083,26300. .26384,
                     27404. .27586,28644. .28804,30189. .30303,33197. .33313,
                     33963. .34128,36231. .36256)
                     /gene="UGCGL1"
                     /note="Homo sapiens UDP-glucose ceramide
                     glucosyltransferase-like 1, mRNA (cDNA clone MGC:46056
                     IMAGE:5492334), complete cds.; H_NH0462M09.1
                     This gene was based on gi(26996809)
                     Continued from H_NH0398G09.1"
                     /codon_start=3
                     /product="unknown"
                     /protein_id="AAY14735.1"
                     /db_xref="GI:62822186"
                     /translation="NFFVDDYARFTILDSQGKTAAVANSMNYLTKKGMSSKEIYDDSF
                     IRPVTFWIVGDFDSPSGRQLLYDAIKHQKSSNNVRISMINNPAKEISYENTQISRAIW
                     AALQTQTSNAAKNFITKMAKEGAAEALAAGADIAEFSVGGMDFSLFKEVFESSKMDFI
                     LSHAVYCRDVLKLKKGQRAVISNGRIIGPLEDSELFNQDDFHLLENIILKTSGQKIKS
                     HIQQLRVEEDVASDLVMKVDALLSAQPKGDPRIEYQFFEDRHSAIKLRPKEGETYFDV
                     VAVVDPVTREAQRLAPLLLVLAQLINMNLRVFMNCQSKLSDMPLKSFYRYVLEPEISF
                     TSDNSFAKGPIAKFLDMPQSPLFTLNLNTPESWMVESVRTPYDLDNIYLEEVDSVVAA
                     EYELEYLLLEGHCYDITTGQPPRGLQFTLGTSANPVIVDTIVMANLGYFQLKANPGAW
                     ILRLRKGRSEDIYRIYSHDGTDSPPDADEVVIVLNNFKSKIIKVKVQKKADMVNEDLL
                     SDGTSENESGFWDSFKWGFTGQKTEEVKQDKDDIINIFSVASGHLYERFLRIMMLSVL
                     KNTKTPVKFWFLKNYLSPTFKEFIPYMANEYNFQYELVQYKWPRWLHQQTEKQRIIWG
                     YKILFLDVLFPLVVDKFLFVDADQIVRTDLKELRDFNLDGAPYGYTPFCDSRREMDGY
                     RFWKSGYWASHLAGRKYHISALYVVDLKKFRKIAAGDRLRGQYQGLSQDPNSLSNLDQ
                     DLPNNMIHQVPIKSLPQEWLWCETWCDDASKKRAKTIDLCNNPMTKEPKLEAAVRIVP
                     EWQDYDQEIKQLQIRFQKEKETGALYKEKTKEPSREGPQKREEL"
     misc_feature    54569. .54786
                     /note="CpG_island (%GC=67.0, o/e=0.76, #CpGs=19)"
     misc_feature    65986. .66196
                     /note="CpG_island (%GC=57.3, o/e=0.94, #CpGs=17)"
     misc_feature    79281. .80265
                     /note="CpG_island (%GC=69.9, o/e=0.83, #CpGs=96)"
     gene            complement(113972. .165192)
                     /gene="HS6ST1"
     mRNA            complement(join(113972. .115384,164551. .165192))
                     /gene="HS6ST1"
     CDS             complement(join(114676. .115384,164551. .165047))
                     /gene="HS6ST1"
                     /note="Homo sapiens heparan sulfate 6-O-sulfotransferase 1
                     (HS6ST1), mRNA.; H_NH0462M09.2
                     This gene was based on gi(4758565)"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAY14736.1"
```

```
                     /db_xref="GI:62822187"
                     /translation="MVERASKFVLVVAGSVCFMLILYQYAGPGLSLGAPGGRAPPDDL
                     DLFPTPDPHYEKKYYFPVRELERSLRFDMKGDDVIVFLHIQKTGGTTFGRHLVQNVRL
                     EVPCDCRPGQKKCTCYRPNRRETWLFSRFSTGWSCGLHADWTELTNCVPGVLDRRDSA
                     ALRTPRKFYYITLLRDPVSRYLSEWRHVQRGATWKTSLHMCDGRTPTPEELPPCYEGT
                     DWSGCTLQEFMDCPYNLANNRQVRMLADLSLVGCYNLSFIPEGKRAQLLLESAKKNLR
                     GMAFPGLTEFQRKTQYLFERTFNLKFIRPFMQYNSTRAGGVEVDEDTIRRIEELNDLD
                     MQLYDYAKDLFQQRYQYKRQLERREQRLRSREERLLHRAKEALPREDADEPGRVPTED
                     YMSHIIEKW"
     misc_feature    114711. .115359
                     /note="CpG_island (%GC=65.8, o/e=0.72, #CpGs=60)"
     misc_feature    164138. .166579
                     /note="CpG_island (%GC=71.7, o/e=0.80, #CpGs=269)"
     misc_feature    168413. .169110
                     /note="CpG_island (%GC=74.4, o/e=0.96, #CpGs=85)"
ORIGIN

  Query Match             36.4%;  Score 130.8;  DB 8;  Length 176075;
  Best Local Similarity   67.2%;  Pred. No. 1.1e-18;
  Matches  215;  Conservative    0;  Mismatches  102;  Indels    3;  Gaps    2;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||||  |||||||||||||| ||| ||||||||||||||||||||
Db      91854 CCCGGCTAATTTTGTAT-GTTTAGTAGAGACGGGGTTTCTCCATGTTGGTCAGGCTGGTC 91796

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              || | || |  ||||||||||||||| ||| ||| ||||||||||||||| |||||||||
Db      91795 TCCAGCTCCTGACCTCAGGTGATCCACCCACCTTGGCCTCCCAAAGTGCTGGGATTACAG 91736

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              || |||||| |||||| | |||    | | |  | | |||  |  |  | |   |  | |
Db      91735 GCATGAGCCTCCGCGCCCGGCCAACAATCTCTATATTTTAATTGGTGTATTTAAACCACT 91676

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATT--TTCT 238
              | |  | |||  ||| |  ||  || | ||| |||| | |  | | | |  ||  |  |
Db      91675 GATGTTCAAAGTGAATATTGATATAGTTAGATTAATATTTACCATATATTTTTTACTATT 91616

Qy        239 TCCCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTTTTCTTTCTAATTTC 298
              | | |  ||    |      |||||| | ||| || || | ||    | | ||||   ||
Db      91615 TTCTATTAGTTGCCCTCGTTCTTCATGCCTATTTTTGCCTTCCACTCTTTTTCTGCCTTT 91556

Qy        299 AAAGGAGTATAATTAAATTG 318
                |||  |  | | | ||||
Db      91555 TGAGGTTTTAATTGAGATTG 91536


RESULT 9
AL137852/c
LOCUS       AL137852               162509 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-131A5 on chromosome 9q32-34.11
            Contains the RAD23B gene for RA23B homolog B (S. cerevisiae) and
            three CpG islands, complete sequence.
ACCESSION   AL137852
VERSION     AL137852.15  GI:13443288
KEYWORDS    HTG; CpG island; HR23B; RAD23B; UV excision repair.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 162509)
  AUTHORS   Beasley,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 24, 2001 this sequence version replaced gi:13276998.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-131A5 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .162509
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /map="q32-34.11"
                     /clone="RP11-131A5"
                     /clone_lib="RPCI-11.1"
     misc_feature    1
                     /note="Clone_left_end: RP11-131A5"
     misc_feature    3638. .3843
                     /note="Sequence from overlapping clone bA417L14
                     (AL445487)"
     misc_feature    42357. .42785
                     /note="Sequence from overlapping clone bA417L14
                     (AL445487)"
     gene            join(44403. .44575,44793. .44961,61407. .61488,63301. .63380,
                     67645. .>67856)
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-002"
     mRNA            join(44403. .44575,44793. .44961,61407. .61488,63301. .63380,
                     67645. .>67856)
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-002"
                     /product="RAD23 homolog B (S. cerevisiae)"
                     /note="match: ESTs: BI561970.1"
     gene            join(44545. .44961,61407. .61488,63301. .63380,67645. .67913,
                     72948. .73003,80018. .80145,83249. .83384,85156. .85283,
                     86100. .86270,90809. .93460)
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-001"
     mRNA            join(44545. .44961,61407. .61488,63301. .63380,67645. .67913,
                     72948. .73003,80018. .80145,83249. .83384,85156. .85283,
                     86100. .86270,90809. .93460)
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-001"
                     /product="RAD23 homolog B (S. cerevisiae)"
                     /note="match: ESTs: AA460535.1 AA460536.1 AI221288.1
                     AL540969.1 AL542437.1 AL569427.1 AU139765.1 AW055317.1
                     AW663949.1 BE018477.1 BE048265.1 BM663704.1 BM977864.1
                     BQ000373.1 BQ002122.1 BQ022617.1 BQ045177.1 BQ772030.1
                     BU508207.1
                     match: cDNAs: D21090.1 X92411.1"
     CDS             join(44896. .44961,61407. .61488,63301. .63380,67645. .67913,
                     72948. .73003,80018. .80145,83249. .83384,85156. .85283,
                     86100. .86270,90809. .90922)
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-001"
                     /standard_name="OTTHUMP00000021857"
                     /note="match: proteins: AAH27747 CAD13275 P54727 P54728
                     Q8WUB0"
                     /codon_start=1
```

```
                     /product="RAD23 homolog B (S. cerevisiae)"
                     /protein_id="CAD13275.1"
                     /db_xref="GI:17384416"
                     /db_xref="Genew:9813"
                     /db_xref="GOA:P54727"
                     /db_xref="InterPro:IPR000449"
                     /db_xref="InterPro:IPR000626"
                     /db_xref="InterPro:IPR004806"
                     /db_xref="InterPro:IPR006636"
                     /db_xref="InterPro:IPR009060"
                     /db_xref="UniProt/Swiss-Prot:P54727"
                     /translation="MQVTLKTLQQQTFKIDIDPEETVKALKEKIESEKGKDAFPVAGQ
                     KLIYAGKILNDDTALKEYKIDEKNFVVVMVTKPKAVSTPAPATTQQSAPASTTAVTSS
                     TTTTVAQAPTPVPALAPTSTPASITPASATASSEPAPASAAKQEKPAEKPAETPVATS
                     PTATDSTSGDSSRSNLFEDATSALVTGQSYENMVTEIMSMGYEREQVIAALRASFNNP
                     DRAVEYLLMGIPGDRESQAVVDPPQAASTGAPQSSAVAAAAATTTATTTTTSSGGHPL
                     EFLRNQPQFQQMRQIIQQNPSLLPALLQQIGRENPQLLQQISQHQEHFIQMLNEPVQE
                     AGGQGGGGGGGSGGIAEAGSGHMNYIQVTPQEKEAIERLKALGFPEGLVIQAYFACEK
                     NENLAANFLLQQNFDED"
CDS                  join(44896. .44961,61407. .61488,63301. .63380,67645. .>67856)
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-002"
                     /standard_name="OTTHUMP00000021858"
                     /codon_start=1
                     /product="RAD23 homolog B (S. cerevisiae)"
                     /protein_id="CAH70394.1"
                     /db_xref="GI:55661955"
                     /db_xref="GOA:Q5W0S5"
                     /db_xref="InterPro:IPR000626"
                     /db_xref="UniProt/TrEMBL:Q5W0S5"
                     /translation="MQVTLKTLQQQTFKIDIDPEETVKALKEKIESEKGKDAFPVAGQ
                     KLIYAGKILNDDTALKEYKIDEKNFVVVMVTKPKAVSTPAPATTQQSAPASTTAVTSS
                     TTTTVAQAPTPVPALAPTSTPASITPASATASSEPAPASAAKQE"
gene                 join(45961. .46179,61407. .61488,63301. .63380,67645. .>67719)
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-003"
mRNA                 join(45961. .46179,61407. .61488,63301. .63380,67645. .>67719)
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-003"
                     /product="RAD23 homolog B (S. cerevisiae)"
                     /note="match: ESTs: AA431131.1"
CDS                  join(46177. .46179,61407. .61488,63301. .63380,67645. .>67719)
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-003"
                     /standard_name="OTTHUMP00000021859"
                     /codon_start=1
                     /product="RAD23 homolog B (S. cerevisiae)"
                     /protein_id="CAH70395.1"
                     /db_xref="GI:55661956"
                     /db_xref="GOA:Q5W0S4"
                     /db_xref="InterPro:IPR000626"
                     /db_xref="UniProt/TrEMBL:Q5W0S4"
                     /translation="MVKALKEKIESEKGKDAFPVAGQKLIYAGKILNDDTALKEYKID
                     EKNFVVVMVTKPKAVSTPAPATTQQSAPASTTAVTS"
gene                 join(<67808. .67913,72948. .73003,80018. .80145,92935. .93455)
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-004"
mRNA                 join(<67808. .67913,72948. .73003,80018. .80145,92935. .93455)
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-004"
                     /product="RAD23 homolog B (S. cerevisiae)"
                     /note="match: ESTs: BQ929186.1"
CDS                  join(<67808. .67913,72948. .73003,80018. .80145,92935. .92988)
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-004"
                     /standard_name="OTTHUMP00000021860"
                     /codon_start=3
                     /product="RAD23 homolog B (S. cerevisiae)"
                     /protein_id="CAH70396.1"
                     /db_xref="GI:55661957"
                     /db_xref="InterPro:IPR000449"
                     /db_xref="UniProt/TrEMBL:Q5W0S3"
                     /translation="TASSEPAPASAAKQEKPAEKPAETPVATSPTATDSTSGDSSRSN
                     LFEDATSALVTGQSYENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLMIIVKTG
                     NKKKKQPLLGK"
misc_feature         84694
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-004"
                     /note="Clone_left_end: RP11-363D24"
misc_feature         84699
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-004"
                     /note="Clone_right_end: RP11-417L14"
misc_feature         87770. .87777
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-004"
                     /note="Single clone region. Sequence from reads from a
                     short insert library derived from a single pUC clone.
                     Restriction digest data confirm the assembly."
polyA_signal         91138. .91143
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-004"
polyA_site           91158
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-004"
polyA_signal         92136. .92141
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-004"
polyA_site           92159
                     /gene="RAD23B"
                     /locus_tag="RP11-131A5.1-004"
polyA_signal         92287. .92292

  Query Match            35.5%;  Score 127.4;  DB 8;  Length 162509;
  Best Local Similarity  89.7%;  Pred. No. 6.2e-18;
  Matches  148;  Conservative    0;  Mismatches   16;  Indels    1;  Gaps    1;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||||| |||||||||| |||| ||| ||||||||||||||||||||||
Db      29321 CCCGGCTAATTTTGTAT-TTTTAGTAGAAACGGGGTTTCTCCATGTTGGTCAGGCTGGTC 29263

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | | |||||||||||||||||||||| |||||||||||||||| |||||||||
Db      29262 TTGAACTCCCAGCCTCAGGTGATCCGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAG 29203

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCT 165
              |||||||||||||| |||||||| ||| |||||||||| ||| |
Db      29202 ACGTGAGCCACCGCGCCCAGCCTGGATACATCTTTTCTTGCATAT 29158


RESULT 10
BD152002
LOCUS       BD152002                 568 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD152002
VERSION     BD152002.1  GI:27857760
KEYWORDS    JP 2002191363-A/6845.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 568)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 6845 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/6845
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
```

```
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                       Location/Qualifiers
            FT   source          1. .568
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .568
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             35.4%;  Score 127;  DB 6;  Length 568;
  Best Local Similarity   79.1%;  Pred. No. 9.2e-18;
  Matches  151;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              || ||||||||||    ||| ||||||||| | ||| ||||||||| ||||||||||||
Db        142 CCTGGCTAATTTTGGTATTTTCAGTAGAGACAGGGTTTCTCCATGTTAGTCAGGCTGGTC 201

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  ||||||| |||||||||| |||||||||||||||||||| |||||||||
Db        202 TCGAACTCCCGACCTCAGATGATCCGCCCACCTCGGCCTCCCAAAGTGCTGGGATTACAG 261

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||| ||| ||||||    | | ||||    | || |||  |||||  | ||
Db        262 GCGTGAGCCACCACGCCCAGCCTAAAGAAATCTTTAAAAATATTTTCTGGTGCTCTACAT 321

Qy        181 GCTTATGAAAA 191
              | | |   |||
Db        322 GTTCAGAGAAA 332


RESULT 11
AX871940
LOCUS       AX871940                 568 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 6845 from Patent EP1074617.
ACCESSION   AX871940
VERSION     AX871940.1  GI:40026767
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 6845 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .568
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             35.4%;  Score 127;  DB 6;  Length 568;
  Best Local Similarity   79.1%;  Pred. No. 9.2e-18;
  Matches  151;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              || ||||||||||    ||| ||||||||| | ||| ||||||||| ||||||||||||
Db        142 CCTGGCTAATTTTGGTATTTTCAGTAGAGACAGGGTTTCTCCATGTTAGTCAGGCTGGTC 201

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  ||||||| |||||||||| |||||||||||||||||||| |||||||||
Db        202 TCGAACTCCCGACCTCAGATGATCCGCCCACCTCGGCCTCCCAAAGTGCTGGGATTACAG 261

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||| ||| ||||||    | | ||||    | || |||  |||||  | ||
Db        262 GCGTGAGCCACCACGCCCAGCCTAAAGAAATCTTTAAAAATATTTTCTGGTGCTCTACAT 321

Qy        181 GCTTATGAAAA 191
              | | |   |||
Db        322 GTTCAGAGAAA 332


RESULT 12
AL512642/c
LOCUS       AL512642               139255 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-469L23 on chromosome 13 Contains
            the ALOX5AP gene for arachidonate 5-lipoxygenase-activating protein
            (FLAP) and a novel gene, complete sequence.
ACCESSION   AL512642
VERSION     AL512642.18  GI:16944077
KEYWORDS    HTG; ALOX5AP; FLAP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 139255)
  AUTHORS   Mashreghi-Mohammadi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 15, 2001 this sequence version replaced gi:16444731.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 13, constructed by the Sanger Centre Chromosome 13
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr13
            RP11-469L23 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .139255
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /clone="RP11-469L23"
                     /clone_lib="RPCI-11.2"
     misc_feature    2000
                     /note="Clone_right_end: RP11-121O19"
     misc_feature    13730. .13735
                     /note="Sequence from uni-directional dGTP big dye
                     terminator reads only."
     gene            join(36755. .36922,45307. .45406,53299. .53369,57191. .57272,
                     65191. .65666)
                     /gene="ALOX5AP"
                     /locus_tag="RP11-469L23.1-001"

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||| ||| ||||||    | | ||||    | || |||  |||||  | ||
Db        262 GCGTGAGCCACCACGCCCAGCCTAAAGAAATCTTTAAAAATATTTTCTGGTGCTCTACAT 321

Qy        181 GCTTATGAAAA 191
              | | |   |||
Db        322 GTTCAGAGAAA 332
```

```
     mRNA            join(36755. .36922,45307. .45406,53299. .53369,57191. .57272,
                     65191. .65666)
                     /gene="ALOX5AP"
                     /locus_tag="RP11-469L23.1-001"
                     /product="arachidonate 5-lipoxygenase-activating protein"
                     /note="match: ESTs: AA822998 AI122464 AV716253 AV744685
                     BF471373 BF796002 BG435742 BG537585 BG548714 BG569867
                     BI911044 BM972232 BQ016999 BQ287903
                     match: cDNAs: BC018538 BC026209 M96552 M96553 M96554
                     M96556 M96557 X52195"
     CDS             join(36853. .36922,45307. .45406,53299. .53369,57191. .57272,
                     65191. .65353)
                     /gene="ALOX5AP"
                     /locus_tag="RP11-469L23.1-001"
                     /standard_name="OTTHUMP00000018204"
                     /note="match: proteins: AAH18538 P20291 P20292 P30353
                     P30354 P30355 P30356 P30357 P30358 Q9D138"
                     /codon_start=1
                     /product="arachidonate 5-lipoxygenase-activating protein"
                     /protein_id="CAH74084.1"
                     /db_xref="GI:55665393"
                     /db_xref="InterPro:IPR001129"
                     /db_xref="InterPro:IPR001446"
                     /translation="MDQETVGNVVLLAIVTLISVVQNGFFAHKVEHESRTQNGRSFQR
                     TGTLAFERVYTANQNCVDAYPTFLAVLWSAGLLCSQVPAAFAGLMYLFVRQKYFVGYL
                     GERTQSTPGYIFGKRIILFLFLMSVAGIFNYYLIFFFGSDFENYIKTISTTISPLLLI
                     P"
     misc_feature    40180. .40242
                     /gene="ALOX5AP"
                     /locus_tag="RP11-469L23.1-001"
                     /note="Single clone region. Assembly confirmed by
                     restriction digest data."
     gene            join(43583. .43792,45307. .45643)
                     /gene="ALOX5AP"
                     /locus_tag="RP11-469L23.1-002"
     mRNA            join(43583. .43792,45307. .45643)
                     /gene="ALOX5AP"
                     /locus_tag="RP11-469L23.1-002"
                     /product="arachidonate 5-lipoxygenase-activating protein"
                     /note="match: ESTs: BF892107 BF892870"
     polyA_signal    65644. .65649
                     /gene="ALOX5AP"
                     /locus_tag="RP11-469L23.1-001"
     polyA_site      65673
     gene            join(104453. .104839,111374. .111892)
                     /locus_tag="RP11-469L23.2-001"
     mRNA            join(104453. .104839,111374. .111892)
                     /locus_tag="RP11-469L23.2-001"
                     /note="match: ESTs: BI818416 BI818490"
     misc_feature    134043. .134186
                     /note="Single clone region. Assembly confirmed by
                     restriction digest data."
     misc_feature    134054. .134063
                     /note="Single clone region. Sequence from uni-directional
                     dGTP big dye terminator reads only."
     misc_feature    139255
                     /note="Clone_right_end: RP11-469L23"
ORIGIN

  Query Match             35.3%;  Score 126.8;  DB 8;  Length 139255;
  Best Local Similarity   84.1%;  Pred. No. 8.4e-18;
  Matches  143;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              | |||| ||||| ||| |||||||||||| || ||| |||||||||||||||||||||||
Db     123819 CTGGCTGATTTTCTATTTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTCT 123760

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              |||||| |  ||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     123759 CGAACTCCTGACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGG 123700

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGT 171
              |||||||||||| ||| |||||  | ||  || |||   | | |||  | |
Db     123699 CGTGAGCCACCACGCCCAGCCAAGCAATTCCATTTTCAAAAGCTTTTAAT 123650


RESULT 13
AL451053
LOCUS       AL451053              148285 bp    DNA     linear   HTG 30-SEP-2001
DEFINITION  Homo sapiens chromosome 1 clone RP11-296P18, 19 unordered pieces.
ACCESSION   AL451053 AC027437
VERSION     AL451053.3  GI:13872392
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hall,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On or before May 15, 2001 this sequence version replaced
            gi:7637297, gi:13785412.
            Draft Sequence Produced by Whitehead Institute/MIT Center for
            Genome Research, 320 Charles Street,
            Cambridge, MA 02141, USA
            http://www-seq.wi.mit.edu
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA296P18
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: M13; M77815; 62% of reads
            Sequencing vector: plasmid; L08752; 37% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 142362 bases at least Q40
            Consensus quality: 144065 bases at least Q30
            Consensus quality: 145103 bases at least Q20
            Insert size: 146485; sum-of-contigs
            Insert size: 163808; 8.2% error; agarose-fp
            Quality coverage: 6.17x in Q20 bases; sum-of-contigs Quality
            coverage: 5.94x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 19 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3047: contig of 3047 bp in length
            *     3048     3147: gap of 100 bp
            *     3148    11710: contig of 8563 bp in length
            *    11711    11810: gap of 100 bp
            *    11811    18527: contig of 6717 bp in length
            *    18528    18627: gap of 100 bp
            *    18628    36547: contig of 17920 bp in length
            *    36548    36647: gap of 100 bp
            *    36648    46245: contig of 9598 bp in length
            *    46246    46345: gap of 100 bp
            *    46346    49690: contig of 3345 bp in length
            *    49691    49790: gap of 100 bp
            *    49791    64661: contig of 14871 bp in length
            *    64662    64761: gap of 100 bp
            *    64762    70030: contig of 5269 bp in length
            *    70031    70130: gap of 100 bp
            *    70131    72597: contig of 2467 bp in length
            *    72598    72697: gap of 100 bp
```

```
            *    72698    76795: contig of 4098 bp in length
            *    76796    76895: gap of 100 bp
            *    76896    84772: contig of 7877 bp in length
            *    84773    84872: gap of 100 bp
            *    84873   106386: contig of 21514 bp in length
            *   106387   106486: gap of 100 bp
            *   106487   113126: contig of 6640 bp in length
            *   113127   113226: gap of 100 bp
            *   113227   122212: contig of 8986 bp in length
            *   122213   122312: gap of 100 bp
            *   122313   124495: contig of 2183 bp in length
            *   124496   124595: gap of 100 bp
            *   124596   128594: contig of 3999 bp in length
            *   128595   128694: gap of 100 bp
            *   128695   132520: contig of 3826 bp in length
            *   132521   132620: gap of 100 bp
            *   132621   141424: contig of 8804 bp in length
            *   141425   141524: gap of 100 bp
            *   141525   148285: contig of 6761 bp in length.
FEATURES             Location/Qualifiers
     source          1. .148285
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-296P18"
                     /clone_lib="RPCI-11.2"
     misc_feature    1. .3047
                     /note="assembly_fragment:01208
                     fragment_chain:1"
     misc_feature    3148. .11710
                     /note="assembly_fragment:00530
                     fragment_chain:1"
     misc_feature    11811. .18527
                     /note="assembly_fragment:00267
                     fragment_chain:1"
     misc_feature    18628. .36547
                     /note="assembly_fragment:01910
                     fragment_chain:2"
     misc_feature    36648. .46245
                     /note="assembly_fragment:02197
                     fragment_chain:2"
     misc_feature    46346. .49690
                     /note="assembly_fragment:00193
                     fragment_chain:2"
     misc_feature    49791. .64661
                     /note="assembly_fragment:00774
                     fragment_chain:3"
     misc_feature    64762. .70030
                     /note="assembly_fragment:00946
                     fragment_chain:3"
     misc_feature    70131. .72597
                     /note="assembly_fragment:02119
                     fragment_chain:4"
     misc_feature    72698. .76795
                     /note="assembly_fragment:01354
                     fragment_chain:4"
     misc_feature    76896. .84772
                     /note="assembly_fragment:00200"
     misc_feature    84873. .106386
                     /note="assembly_fragment:00242"
     misc_feature    106487. .113126
                     /note="assembly_fragment:00635"
     misc_feature    113227. .122212
                     /note="assembly_fragment:00884"
     misc_feature    122313. .124495
                     /note="assembly_fragment:01118"
     misc_feature    124596. .128594
                     /note="assembly_fragment:01225"
     misc_feature    128695. .132520
                     /note="assembly_fragment:01240"
     misc_feature    132621. .141424
                     /note="assembly_fragment:01260"
     misc_feature    141525. .148285
                     /note="assembly_fragment:01438"
ORIGIN

  Query Match             35.3%;  Score 126.8;  DB 14;  Length 148285;
  Best Local Similarity   76.7%;  Pred. No. 8.4e-18;
  Matches  155;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              || ||||||||||||||| |||||||||||||| | ||| |||||||||||| ||||||||||
Db     128346 CCTGGCTAATTTTGTATTTTTTAGTAGAGACAGGGTTTCTCCATGTTGGCCAGGCTGGTC 128405

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| ||||||||||||||||||||||| |||||||||
Db     128406 TCGAACTCCCGACCTCAGGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 128465

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              || |||||||||||||| | ||||   | | |  ||||     | | |   | |   ||||
Db     128466 GCATGAGCCACCGCGCCCGGCCTAACATCTCTATTTCAACAGTATACCTTTTCCTAAAAT 128525

Qy        181 GCTTATGAAAACGAAAAAAGAA 202
              | | ||     ||| | | |
Db     128526 GTATTTGCCTGTGAATTAGGGA 128547


RESULT 14
AL590133/c
LOCUS       AL590133               192096 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-316M1 on chromosome 1 Contains
            the SETDB1 gene for SET domain, bifurcated 1, the LASS2 gene for
            LAG1 longevity assurance homolog 2 (S. cerevisiae), the ANXA9 gene
            for annexin A9, the gene for a novel protein (FLJ11280), the gene
            for TcD37 homolog (HTCD37), the BNIPL gene for BCL2/adenovirus E1B
            19kD interacting protein like, the gene for a novel protein
            (FLJ20519), the gene for small protein effector 1 of Cdc42 (SPEC1),
            the gene for ALL1-fused gene from chromosome 1q (AF1Q), two novel
            genes, the 5' end of a novel gene (MGC29891) and two CpG islands,
            complete sequence.
ACCESSION   AL590133 AC073204
VERSION     AL590133.32  GI:20218562
KEYWORDS    HTG; AF1Q; ANXA9; BNIPL; CDC42; CpG island; FLJ11280; FLJ20519;
            HTCD37; LASS2; MGC29891; SETDB1; SPEC1; TcD37.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 192096)
  AUTHORS   Bird,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 19, 2002 this sequence version replaced gi:20196555.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
```

```
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
            RP11-316M1 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            Draft Sequence Produced by Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            http://genome.wustl.edu/gsc/index.shtml.
FEATURES              Location/Qualifiers
     source           1. .192096
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="1"
                      /clone="RP11-316M1"
                      /clone_lib="RPCI-11.2"
     misc_feature     1
                      /note="Clone_left_end: RP11-316M1"
     gene             join(12484. .12541,13897. .14167,16160. .16230)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-006"
     mRNA             join(12484. .12541,13897. .14167,16160. .16230)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-006"
                      /product="SET domain, bifurcated 1"
                      /note="match: ESTs: Em:BX472886.1 Em:BX472895.1"
     gene             join(12605. .12710,13897. .14167,16160. .16311,26133. .26167,
                      27522. .27621,28756. .28881,29045. .29246,30113. .30186,
                      31111. .31301,33079. .33205,35315. .35471,35563. .35718,
                      36651. .37283,37558. .37674,45371. .45540,46756. .47384,
                      48323. .48351,48780. .48912,49167. .49332,49723. .49934,
                      50188. .50276,50440. .50930)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-001"
     mRNA             join(12605. .12710,13897. .14167,16160. .16311,26133. .26167,
                      27522. .27621,28756. .28881,29045. .29246,30113. .30186,
                      31111. .31301,33079. .33205,35315. .35471,35563. .35718,
                      36651. .37283,37558. .37674,45371. .45540,46756. .47384,
                      48323. .48351,48780. .48912,49167. .49332,49723. .49934,
                      50188. .50276,50440. .50930)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-001"
                      /product="SET domain, bifurcated 1"
                      /note="match: ESTs: Em:AI499881.1 Em:AI582266.1
                      Em:AI632489.1 Em:AI659916.1 Em:AI830554.1 Em:AU151491.1
                      Em:AW351749.1 Em:AW451316.1 Em:BE256325.1 Em:BE274369.1
                      Em:BE395669.1 Em:BE544349.1 Em:BE727922.1 Em:BE727991.1
                      Em:BE784129.1 Em:BE790566.1 Em:BE792839.1 Em:BF026775.1
                      Em:BF241505.1 Em:BG331938.1 Em:BG333121.1 Em:BG333851.1
                      Em:BG342188.1 Em:BG437451.1 Em:BG685286.1 Em:BG745393.1
                      Em:BG753913.1 Em:BG755888.1 Em:BM011729.1 Em:BM464383.1
                      Em:BM558159.1 Em:BQ675487.1 Em:BQ917875.1 Em:BQ935376.1
                      Em:BQ956716.1 Em:BU500506.1 Em:BU624516.1 Em:BU682854.1
                      Em:CB153628.1 Em:CB270416.1 Em:CB306095.1 Em:CF136719.1
                      match: cDNAs: Em:BC028671.2 Em:D31891.1"
     gene             join(12617. .12656,13897. .14167,16160. .16311,46756. .47083)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-003"
     mRNA             join(12617. .12656,13897. .14167,16160. .16311,46756. .47083)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-003"
                      /product="SET domain, bifurcated 1"
                      /note="match: ESTs: Em:BQ220240.1"
     gene             join(12617. .12710,13897. .14167,16160. .16311,26133. .26167,
                      27522. .27621,28756. .28881,30113. .30186,31111. .31303)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-009"
     mRNA             join(12617. .12710,13897. .14167,16160. .16311,26133. .26167,
                      27522. .27621,28756. .28881,30113. .30186,31111. .31303)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-009"
                      /product="SET domain, bifurcated 1"
                      /note="match: ESTs: Em:BX440371.1"
     gene             join(12642. .12656,13897. .14167,16160. .16311,26133. .26167,
                      27522. .27621,28756. .28881,29045. .29246,30113. .30186,
                      31111. .31513)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-002"
     mRNA             join(12642. .12656,13897. .14167,16160. .16311,26133. .26167,
                      27522. .27621,28756. .28881,29045. .29246,30113. .30186,
                      31111. .31513)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-002"
                      /product="SET domain, bifurcated 1"
                      /note="match: ESTs: Em:AA576434.1 Em:AI160504.1
                      Em:AI377943.1 Em:AI762193.1 Em:AW235046.1 Em:AW292360.1
                      Em:BI855763.1 Em:BQ004073.1 Em:BQ055551.1 Em:BU621595.1
                      Em:BU940444.1 Em:CA422418.1 Em:CA444580.1
                      match: cDNAs: Em:BC009362.1 Em:BC009362.2"
     gene             join(12833. .13001,13897. .14167,16160. .>16309)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-008"
     mRNA             join(12833. .13001,13897. .14167,16160. .>16309)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-008"
                      /product="SET domain, bifurcated 1"
                      /note="match: ESTs: Em:CB149707.1"
     CDS              join(12971. .13001,13897. .14167,16160. .>16309)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-008"
                      /standard_name="OTTHUMP00000037714"
                      /codon_start=1
                      /product="SET domain, bifurcated 1"
                      /protein_id="CAI13325.1"
                      /db_xref="GI:55960631"
                      /db_xref="UniProt/TrEMBL:Q5SZD9"
                      /translation="MCEVFGLWILEDKSMSSLPGCIGLDAATATVESEEIAELQQAVV
                      EELGISMEELRHFIDEELEKMDCVQQRKKQLAELETWVIQKESEVAHVDQLFDDASRA
                      VTNCESLVKDFYSKLGLQYRDSSSEDESSRPTEIIEIPDEDDDVLSIDS"
     CDS              join(13908. .14167,16160. .16311,26133. .26167,27522. .27621,
                      28756. .28881,29045. .29246,30113. .30186,31111. .31301,
                      33079. .33205,35315. .35471,35563. .35718,36651. .37283,
                      37558. .37674,45371. .45540,46756. .47384,48323. .48351,
                      48780. .48912,49167. .49332,49723. .49934,50188. .50276,
                      50440. .50557)
                      /gene="SETDB1"
                      /locus_tag="RP11-316M1.1-001"
                      /standard_name="OTTHUMP00000032929"
                      /note="match: proteins: Sw:O88974 Sw:Q15047 Tr:Q8BTV6"
                      /codon_start=1
                      /product="SET domain, bifurcated 1"
                      /protein_id="CAI13328.1"
                      /db_xref="GI:55960634"
                      /db_xref="InterPro:IPR001214"
                      /db_xref="InterPro:IPR001739"
                      /db_xref="InterPro:IPR002999"
                      /db_xref="InterPro:IPR003606"
                      /db_xref="InterPro:IPR003616"
                      /db_xref="InterPro:IPR007728"
                      /translation="MSSLPGCIGLDAATATVESEEIAELQQAVVEELGISMEELRHFI
                      DEELEKMDCVQQRKKQLAELETWVIQKESEVAHVDQLFDDASRAVTNCESLVKDFYSK
                      LGLQYRDSSSEDESSRPTEIIEIPDEDDDVLSIDSGDAGSRTPKDQKLREAMAALRKS
                      AQDVQKFMDAVNKKSSSQDLHKGTLSQMSGELSKDGDLIVSMRILGKKRTKTWHKGTL
                      IAIQTVGPGKKYKVKFDNKGKSLLSGNHIAYDYHPPADKLYVGSRVVAKYKDGNQVWL
                      YAGIVAETPNVKNKLRFLIFFDDGYASYVTQSELYPICRPLKKTWEDIEDISCRDFIE
                      EYVTAYPNRPMVLLKSGQLIKTEWEGTWWKSRVEEVDGSLVRILFLDDKRCEWIYRGS
                      TRLEPMFSMKTSSASALEKKQGQLRTRPNMGAVRSKGPVVQYTQDLTGTGTQFKPVEP
                      PQPTAPPAPPFPPAPPLSPQAGDSDLESQLAQSRKQVAKKSTSFRPGSVGSGHSSPTS
                      PALSENVSGGKPGINQTYRSPLGSTASAPAPSALPAPPAPPVFHGMLERAPAEPSYRA
                      PMEKLFYLPHVCSYTCLSRVRPMRNEQYRGKNPLLVPLLYDFRRMTARRVNRKMGFH
                      VIYKTPCGLCLRTMQEIERYLFETGCDFLFLEMFCLDPYVLVDRKFQPYKPFYYILDI
                      TYGKEDVPLSCVNEIDTTPPPQVAYSKERIPGKGVFINTGPEFLVGCDCKDGCRDKSK
                      CACHQLTIQATACTPGGQINPNSGYQYKRLEECLPTGVYECNKRCKCDPNMCTNRLVQ
```

```
                     HGLQVRLQLFKTQNKGWGIRCLDDIAKGSFVCIYAGKILTDDFADKEGLEMGDEYFAN
                     LDHIESVENFKEGYESDAPCSSDSSGVDLKDQEDGNSGTEDPEESNDDSSDDNFCKDE
                     DFSTSSVWRSYATRRQTRGQKENGLSETTSKDSHPPDLGPPHIPVPPSIPVGGCNPPS
                     SEETPKNKVASWLSCNSVSEGGFADSDSHSSFKTNEGGEGRAGGSRMEAEKASTSGLG
                     IKDEGDIKQAKKEDTDDRNKMSVVTESSRNYGYNPSPVKPEGLRRPPSKTSMHQSRRL
                     MASAQSNPDDVLTLSSSTESEGESGTSRKPTAGQTSATAVDSDDIQTISSGSEGDDFE
                     DKKNMTGPMKRQVAVKSTRGFALKSTHGIAIKSTNMASVDKGESAPVRKNTRQFYDGE
                     ESCYIIDAKLEGNLGRYLNHSCSPNLFVQNVFVDTHDLRFPWVAFFASKRIRAGTELT

  Query Match             35.3%;  Score 126.8;  DB 8;  Length 192096;
  Best Local Similarity   76.7%;  Pred. No. 8.3e-18;
  Matches  155;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              || ||||||||||||||| |||||||||||||| | ||| |||||||||||| ||||||||||
Db      36286 CCTGGCTAATTTTGTATTTTTTAGTAGAGACAGGGTTTCTCCATGTTGGCCAGGCTGGTC 36227

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| ||||||||||||||||||||||| |||||||||
Db      36226 TCGAACTCCCGACCTCAGGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 36167

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              || |||||||||||||| | ||||   | | |  ||||     | | |   | |   ||||
Db      36166 GCATGAGCCACCGCGCCCGGCCTAACATCTCTATTTCAACAGTATACCTTTTCCTAAAAT 36107

Qy        181 GCTTATGAAAACGAAAAAAGAA 202
              |  | ||     |||  | | |
Db      36106 GTATTTGCCTGTGAATTATGGA 36085


RESULT 15
AC117378/c
LOCUS       AC117378              110459 bp    DNA     linear   PRI 15-MAR-2003
DEFINITION  Homo sapiens 12 BAC RP11-112B10 (Roswell Park Cancer Institute
            Human BAC Library) complete sequence.
ACCESSION   AC117378 AC021586
VERSION     AC117378.8  GI:23346656
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 110459)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S.,
            Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
            Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,
            Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P.,
            Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M.,
            Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S.,
            Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
            Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U.,
            King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N.,
            Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O.,
            Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X.,
            Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P.,
            Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
            Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S.,
            Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
            Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
            Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 110459)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 110459)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 110459)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   5  (bases 1 to 110459)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   6  (bases 1 to 110459)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAR-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 30, 2002 this sequence version replaced gi:21490136.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
```

a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

```
FEATURES             Location/Qualifiers
     source          1. .110459
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /clone="RP11-112B10"
     misc_feature    1. .2002
                     /note="overlaps bases 91579. .93582 of clone AC090681"
                     /function="clone overlap"
     repeat_region   complement(538. .637)
                     /rpt_family="AluY"
     repeat_region   797. .984
                     /rpt_family="AluJb"
     repeat_region   complement(985. .1277)
                     /rpt_family="AluSx"
     repeat_region   1278. .1301
                     /rpt_family="AluJb"
     repeat_region   complement(1306. .1441)
                     /rpt_family="AluJb"
     repeat_region   1442. .1693
                     /rpt_family="AluSq"
     repeat_region   complement(1694. .1799)
                     /rpt_family="AluJb"
     repeat_region   1974. .2086
                     /rpt_family="L2"
     repeat_region   complement(4121. .4390)
                     /rpt_family="AluSg"
     repeat_region   complement(4484. .4756)
                     /rpt_family="AluSx"
     repeat_region   complement(5469. .5762)
                     /rpt_family="AluSx"
     repeat_region   complement(5763. .6055)
                     /rpt_family="AluSx"
     repeat_region   6272. .6661
                     /rpt_family="L1MB1"
     repeat_region   6662. .6959
                     /rpt_family="AluSq"
     repeat_region   6960. .7005
                     /rpt_family="L1MB1"
     STS             7077. .7300
                     /standard_name="RH45308"
     repeat_region   7268. .7396
                     /rpt_family="AluJb"
     repeat_region   complement(7844. .8193)
                     /rpt_family="AluJb"
     repeat_region   complement(8205. .8418)
                     /rpt_family="AluSq"
     repeat_region   complement(8454. .8744)
                     /rpt_family="AluSc"
     repeat_region   8803. .8824
                     /rpt_family="AT_rich"
     repeat_region   complement(9270. .9582)
                     /rpt_family="7SLRNA"
     repeat_region   9665. .9948
                     /rpt_family="AluJo"
     repeat_region   9964. .10247
                     /rpt_family="AluJb"
     repeat_region   10862. .10883
                     /rpt_family="AT_rich"
     repeat_region   11058. .11351
                     /rpt_family="AluSp"
     repeat_region   12184. .12347
                     /rpt_family="AluYb8"
     repeat_region   12348. .12649
                     /rpt_family="AluSx"
     repeat_region   12650. .12823
                     /rpt_family="AluYb8"
     repeat_region   12920. .13245
                     /rpt_family="MLT1B"
     repeat_region   13246. .13383
                     /rpt_family="AluSp"
     repeat_region   13384. .13690
                     /rpt_family="AluY"
     repeat_region   13691. .13863
                     /rpt_family="AluSp"
```

```
Query Match             35.1%;  Score 126;  DB 8;  Length 110459;
Best Local Similarity   84.9%;  Pred. No. 1.3e-17;
Matches  141;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||||| | |||||||||||| || ||| ||||||||||||||||||||
Db      29694 CCCGGCTAATTTTGTGTTTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTC 29635

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| | |||||||||||||| |||||||||||||||||||||| |||| |||||||||
Db      29634 TCGAACTCCCAACCTCAGGTGATTCGCCCGCCTCGGCCTCCCAAAATGCTGGGATTACAG 29575

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTT 166
              ||||||||||||||| | ||||       | ||| |  ||| ||
Db      29574 GCGTGAGCCACCGCGCCCGGCCTCCAGGAGACATTTTTACCATTTT 29529
```

Search completed: May 6, 2006, 20:09:23
Job time : 2420 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: May 5, 2006, 17:29:32 ; Search time 401 Seconds
(without alignments)
5966.645 Million cell updates/sec

Title: US-10-009-579A-5_COPY_3188_3546
Perfect score: 359
Sequence: 1 cccggctaattttgtatctt.....ttttttatagtgttctggaa 359

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	359	100.0	4282	4	AAL41898	Aal41898 Human GA7
	2	127	35.4	568	4	AAH10010	Aah10010 Human cDN
c	3	126.8	35.3	98800	12	ADN06353_3	Continuation (4 of
c	4	126.8	35.3	98800	13	ADS94372_3	Continuation (4 of
	5	124.4	34.7	1744	6	ABZ70301	Abz70301 Human tyr
c	6	123.8	34.5	344	4	AAK68705	Aak68705 Human imm
	7	123.8	34.5	438	5	ABV13839	Abv13839 Human pro
	8	122.4	34.1	160361	12	ADL08116	Adl08116 Human gen
c	9	122.2	34.0	22651	4	AAK78202	Aak78202 Human imm
c	10	122	34.0	313001	14	ADZ70075	Adz70075 Human ins
c	11	122	34.0	321019	13	ADS36450	Ads36450 Human aut
c	12	122	34.0	329019	13	ABD32707	Abd32707 Human can
c	13	121.8	33.9	381	6	ABL83966	Abl83966 Human ova
c	14	121.8	33.9	546	13	ACF91521	Acf91521 Human SIR
c	15	121.8	33.9	556	5	AAS90976	Aas90976 DNA encod
	16	121.8	33.9	3470	5	ABV25066	Abv25066 Human pro
	17	121.8	33.9	122673	14	AEA61123	Aea61123 Human PDE
c	18	121.4	33.8	30393	4	AAK67239	Aak67239 Human imm
c	19	121.2	33.8	291	11	ADZ57848	Adz57848 Human Alu
	20	121.2	33.8	7793	8	ABZ73859	Abz73859 Secreted
	21	121.2	33.8	7793	8	ADA98513	Ada98513 Human sec
	22	121.2	33.8	7793	8	ADA44266	Ada44266 Human sec
	23	121.2	33.8	7793	10	ADC20639	Adc20639 Human sec
	24	121.2	33.8	7793	10	ADF10839	Adf10839 Human sec
	25	121.2	33.8	7793	10	ABZ67436	Abz67436 Human sec
c	26	120.8	33.6	779	4	AAK90274	Aak90274 Human dig
c	27	120.8	33.6	779	4	AAI57654	Aai57654 Human col
c	28	120.8	33.6	779	6	ABS99831	Abs99831 Genomic D
c	29	120.8	33.6	779	10	ADB92984	Adb92984 Human col
c	30	120.8	33.6	1559	12	ADO15908	Ado15908 4 synthes
c	31	120.8	33.6	10159	4	AAK73470	Aak73470 Human imm
c	32	120.8	33.6	10159	4	AAK73471	Aak73471 Human imm
c	33	120.8	33.6	67253	14	AEA61178	Aea61178 Human GPR
c	34	120.4	33.5	47188	10	ADL13758	Adl13758 Osteoarth
c	35	120.2	33.5	3062	4	AAH14445	Aah14445 Human cDN
c	36	120.2	33.5	3062	6	ABX04191	Abx04191 Human mRN
c	37	120.2	33.5	4088	4	AAK87132	Aak87132 Human imm
c	38	120.2	33.5	83517	13	ABD32581	Abd32581 Human can
c	39	120	33.4	301	10	ADH59595	Adh59595 Alu-repea
	40	120	33.4	2744	4	AAK79905	Aak79905 Human imm
	41	120	33.4	2744	4	AAK79904	Aak79904 Human imm
c	42	120	33.4	11581	14	AEA61110	Aea61110 Human CDA
	43	120	33.4	110000	12	ADN06353_0	Adn06353 Human FLA
	44	120	33.4	110000	13	ADS94372_0	Ads94372 Human 5-1
	45	120	33.4	165199	6	ABK83460	Abk83460 Human cDN

ALIGNMENTS

RESULT 1
AAL41898
ID AAL41898 standard; DNA; 4282 BP.
XX
AC AAL41898;
XX
DT 03-MAY-2002 (first entry)
XX
DE Human GA733-2 gene (encoding human epithelial glycoprotein-2) promoter.
XX
KW Human; GA733-2 gene promoter; gene; epithelial glycoprotein-2; EGP-2;
KW pan-carcinoma associated antigen; cancer; carcinoma selective expression;
KW treatment evaluation; non-squamous epithelium disease; carcinogenesis;
KW transgenic animal; ds; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1130106-A1.
XX
PD 05-SEP-2001.
XX
PF 01-MAR-2000; 2000EP-00200728.
XX
PR 01-MAR-2000; 2000EP-00200728.
XX
PA (UYGR-) RIJKSUNIV GRONINGEN.
XX
PI De Leij LFMH, Mclaughlin PMJ, Ruiters MHJ, Harmsen MC;
PI Van Der Molen H, Terpstra P, Dokter WHA;
XX
DR WPI; 2001-591523/67.
XX
PT Novel isolated and/or recombinant nucleic acid having tissue specific
PT promoter derived from epithelial glycoprotein 2 gene, that allows
PT expression of desired nucleic acid in cancer cell, specifically carcinoma
PT cell.
XX
PS Claim 3; Fig 1; 21pp; English.
XX
CC The invention comprises the promoter sequence of the human GA733-2 gene.
CC The GA733-2 gene encodes human epithelial glycoprotein-2 (EGP-2), which
CC is a pan-carcinoma associated antigen. The GA733-2 gene promoter allows

```
CC   the expression of a nucleic acid of interest in a cancer cell, where the
CC   expression within the cancer cell is carcinoma selective. The GA733-2
CC   gene promoter sequence is useful in the treatment of cancer and in
CC   evaluating the treatment of a disease (e.g. a disease of the non-squamous
CC   epithelium, such as carcinogenesis). The GA733-2 gene promoter can also
CC   be used in the production of host cells and a transgenic animal. The
CC   present nucleotide sequence represents the GA733-2 gene promoter sequence
CC   of the invention
XX
SQ   Sequence 4282 BP; 1115 A; 1031 C; 976 G; 1160 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 359;  DB 4;  Length 4282;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-70;
  Matches  359;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3188 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 3247

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3248 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 3307

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3308 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 3367

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3368 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 3427

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTTTTCTTTCTAATTTCAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3428 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTTTTCTTTCTAATTTCAA 3487

Qy        301 AGGAGTATAATTAAATTGCCAGGTAAAAGCTCAAAGGTCTTTTTTATAGTGTTCTGGAA 359
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3488 AGGAGTATAATTAAATTGCCAGGTAAAAGCTCAAAGGTCTTTTTTATAGTGTTCTGGAA 3546


RESULT 2
AAH10010
ID   AAH10010 standard; cDNA; 568 BP.
XX
AC   AAH10010;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (3'-primer) SEQ ID NO:6845.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 3; SEQ ID NO 6845; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 568 BP; 154 A; 138 C; 128 G; 141 T; 0 U; 7 Other;

  Query Match              35.4%;  Score 127;  DB 4;  Length 568;
  Best Local Similarity    79.1%;  Pred. No. 1e-18;
  Matches  151;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              || |||||||||||     ||| ||||||||| | ||| ||||||||| ||||||||||||
Db        142 CCTGGCTAATTTTGGTATTTTCAGTAGAGACAGGGTTTCTCCATGTTAGTCAGGCTGGTC 201

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              |||||||  |  ||||||| |||||||||| |||||||||||||||||||| |||||||||
Db        202 TCGAACTCCCGACCTCAGATGATCCGCCCACCTCGGCCTCCCAAAGTGCTGGGATTACAG 261

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||||||| ||| ||||||     | | ||||    | || |||  |||||  | ||
Db        262 GCGTGAGCCACCACGCCCAGCCTAAAGAAATCTTTAAAAATATTTTCTGGTGCTCTACAT 321

Qy        181 GCTTATGAAAA 191
              | | |   |||
Db        322 GTTCAGAGAAA 332


RESULT 3
ADN06353_3/c
Continuation (4 of 4) of ADN06353 from base 300001 (Human FLAP genomic DNA SEQ ID NO:1.
WP Sequence split into 4 fragments  LOCUS ADN06353 Accession Adn06353
WP      Fragment Name      Begin      End
WP      ADN06353_0             1    110000
WP      ADN06353_1        100001    210000
WP      ADN06353_2        200001    310000
WP      ADN06353_3        300001    398800

  Query Match              35.3%;  Score 126.8;  DB 12;  Length 98800;
  Best Local Similarity    84.1%;  Pred. No. 2.9e-18;
  Matches  143;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              | |||| ||||| ||| |||||||||||| || ||| ||||||||||||||||||||||
Db      81300 CTGGCTGATTTTCTATTTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTCT 81241

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              |||||| |  |||||||||||||||||||||||||||||||||||||||||| ||||||||||
```

```
Db      81240 CGAACTCCTGACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGG 81181

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGT 171
              |||||||||||| ||| |||||  | ||  || |||   | | |||  | |
Db      81180 CGTGAGCCACCACGCCCAGCCAAGCAATTCCATTTTCAAAAGCTTTTAAT 81131


RESULT 4
ADS94372_3/c
Continuation (4 of 4) of ADS94372 from base 300001 (Human 5-lipoxygenase activating prot
WP Sequence split into 4 fragments  LOCUS ADS94372 Accession Ads94372
WP      Fragment Name      Begin     End
WP      ADS94372_0             1     110000
WP      ADS94372_1        100001     210000
WP      ADS94372_2        200001     310000
WP      ADS94372_3        300001     398800

  Query Match             35.3%;  Score 126.8;  DB 13;  Length 98800;
  Best Local Similarity   84.1%;  Pred. No. 2.9e-18;
  Matches  143;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              | |||| ||||| ||| |||||||||||| || ||| ||||||||||||||||||||||
Db      81300 CTGGCTGATTTTCTATTTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTCT 81241

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              ||||||  |  ||||||||||||||||||||||||||||||||||||||||| |||||||||
Db      81240 CGAACTCCTGACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGG 81181

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGT 171
              |||||||||||| ||| |||||  | ||  || |||   | | |||  | |
Db      81180 CGTGAGCCACCACGCCCAGCCAAGCAATTCCATTTTCAAAAGCTTTTAAT 81131


RESULT 5
ABZ70301
ID   ABZ70301 standard; cDNA; 1744 BP.
XX
AC   ABZ70301;
XX
DT   25-APR-2003  (first entry)
XX
DE   Human tyrosinase 10.01 coding sequence.
XX
KW   Human; tyrosinase 10.01; enzyme; cancer; HIV infection; cytostatic;
KW   anti-HIV; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             129. .404
FT                   /*tag=  a
FT                   /product= "Tyrosinase 10.01"
XX
PN   CN1363662-A.
XX
PD   14-AUG-2002.
XX
PF   05-JAN-2001; 2001CN-00105035.
XX
PR   05-JAN-2001; 2001CN-00105035.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2002-751776/82.
DR   P-PSDB; ABP59183.
XX
PT   Polypeptide-human tyrosinase 10.01 and polynucleotide for coding it.
XX
PS   Claim 6; Page 24-25 (Disclosure); 32pp; Chinese.
XX
CC   The present sequence is the coding sequence for human tyrosinase 10.01.
CC   The protein can be used for treating diseases such as cancer and HIV
CC   infection
XX
SQ   Sequence 1744 BP; 354 A; 478 C; 476 G; 436 T; 0 U; 0 Other;

  Query Match             34.7%;  Score 124.4;  DB 6;  Length 1744;
  Best Local Similarity   86.7%;  Pred. No. 4.7e-18;
  Matches  137;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||      |||||||||||||||| ||| ||||||||||||||||||||||
Db       1384 CCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCTCCATGTTGGTCAGGCTGGTC 1443

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              |||||||  |  |||||| |||||||||||||||||||||| |||||||||||| |||||||||
Db       1444 TCGAACTCCTGACCTCATGTGATCCGCCCGCCTCGGCTTCCCAAAGTGCTGGGATTACAG 1503

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCT 158
              ||||||||  ||||||| ||||||| | | | | ||| |
Db       1504 GCGTGAGCTACCGCGCCCAGCCTGCGTAGAGCCTTTTT 1541


RESULT 6
AAK68705/c
ID   AAK68705 standard; DNA; 344 BP.
XX
AC   AAK68705;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23517.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
```

PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 23517; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 344 BP; 91 A; 93 C; 90 G; 70 T; 0 U; 0 Other;

Query Match 34.5%; Score 123.8; DB 4; Length 344;
Best Local Similarity 83.8%; Pred. No. 4.8e-18;
Matches 140; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
||| ||||||||| |||||||||||| || ||| ||||||||||||||||||||
Db 202 CCCAGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTC 143

```
Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| || |||||||||||||||||||||| | |||||||||||| |||||||||
Db        142 TCGAACTCCAGACCTCAGGTGATCCGCCCGCCTCAGTCTCCCAAAGTGCTGGGATTACAG 83

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTC 167
              |||||||||||| |||| | |||||  || || ||||   |||   ||
Db         82 GCGTGAGCCACTGCGCCCGGCCTGATAAAACTTTTTAAAACAGGCTC 36


RESULT 7
ABV13839
ID   ABV13839 standard; cDNA; 438 BP.
XX
AC   ABV13839;
XX
DT   13-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 13830.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005171.
XX
PR   17-FEB-2000; 2000US-0183319P.
PR   16-MAR-2000; 2000US-0189862P.
PR   25-MAY-2000; 2000US-0207454P.
PR   09-JUN-2000; 2000US-0211314P.
PR   18-JUL-2000; 2000US-0219007P.
PR   13-DEC-2000; 2000US-0255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS   Claim 1; Page 2303; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for: (a) assessing whether
CC   a patient is afflicted with prostate cancer; (b) monitoring the
CC   progression of prostate cancer in a patient; (c) assessing the efficacy
CC   of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC   the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC   determining whether prostate cancer has metastasized in a patient; (h)
CC   assessing the aggressiveness or indolence of prostate cancer in a patient
CC   ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ   Sequence 438 BP; 110 A; 102 C; 83 G; 143 T; 0 U; 0 Other;

  Query Match             34.5%;  Score 123.8;  DB 5;  Length 438;
  Best Local Similarity   68.8%;  Pred. No. 5e-18;
  Matches  170;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              ||||||||||||     |||||||||||| || ||| | ||||||||| ||||||| |||
Db        109 CCGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGATCT 168

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              | |||| |  |||||| ||| ||| ||| ||| |||||||||||||||| ||||||||||
Db        169 CAAACTCCTGACCTCATGTGGTCCACCCACCTTGGCCTCCCAAAGTGCTGGGATTACAGG 228

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAATG 181
              ||||||||||||||||||||||    | ||  ||||   | |    ||  |  ||   ||
Db        229 CGTGAGCCACCGCGCTCAGCCTACAAGCAATTTTTTAAAAACTGACACCTATTACTGATA 288

Qy        182 CTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTCC 241
                  |      |||||    |||| |   |  ||| |||  ||  | | |||||  ||
Db        289 AAATTCTCTGTTTAAAAATTCCTATTTACTTTTTTTAAAAAAAAGTATTAATTTTTTTTT 348

Qy        242 CAAGAGA 248
              || | |
Db        349 AAATACA 355


RESULT 8
ADL08116
ID   ADL08116 standard; DNA; 160361 BP.
XX
AC   ADL08116;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human gene associated with low HDL-C FABP-3.
XX
KW   Human; ds; SNP; single nucleotide polymorphism;
KW   high density lipoprotein-C; HDL-C; vascular disease; metabolic disease;
KW   coronary artery disease; gene.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(120178,C)
FT                   /*tag=  a
FT                   /standard_name= "Single nucleotide polymorphism"
XX
PN   US2004043389-A1.
XX
PD   04-MAR-2004.
XX
PF   04-SEP-2002; 2002US-00235192.
XX
PR   04-SEP-2002; 2002US-00235192.
XX
PA   (VITI-) VITIVITY INC.
XX
PI   Mccarthy J;
XX
DR   WPI; 2004-214170/20.
XX
PT   Determining whether a subject has, or is at risk of developing, an
PT   abnormally low high density lipoprotein-C (HDL-C) level comprises
PT   detecting an allelic variant of a polymorphic region from any of a set of
PT   27 genes.
XX
PS   Disclosure; SEQ ID NO 35; 37pp; English.
XX
CC   The invention relates to determining whether a subject has, or is at risk
CC   of developing, an abnormally low high density lipoprotein-C (HDL-C) level
CC   comprises determining whether the subject has an allelic variant of a
CC   polymorphic region from any of 27 genes (alleles listed in Table 5 of the
CC   specification). Also included are determining whether a male subject has,
CC   or is at risk of developing, an abnormally low HDL-C level, comprising
CC   determining whether the male subject has an allelic variant of a
CC   polymorphic region listed in Table 5 which is associated with abnormally
CC   low HDL-C levels in males, and determining whether a female subject has,
CC   or is at risk of developing, an abnormally low HDL-C level, comprising
CC   determining whether the female subject has an allelic variant of a
CC   polymorphic region listed in Table 5 which is associated with abnormally
CC   low HDL-C levels in females. The allelic variant in determining whether a
CC   subject has, or is at risk of developing, an abnormally low HDL-C level
```

```
CC   is APOA 1 CC, CD14 1 CT, COL5A2 1 GG, EDNRB 1 AG or AA, FABP3 1 CT, GBE1
CC   1 AG or GG, LIPC 5 AA, MTHFR 1 CC, VWF 2 GG, or their complements. The
CC   allelic variant in determining whether a male subject has, or is at risk
CC   of developing, an abnormally low HDL-C level, LRP1 3 CC or CT, PAI2 4 GG,
CC   or PPARG 1 CG, or their complements. The allelic variants are also COL5A2
CC   1 GG, CD14 1 CT or CC, and FABP3 1 CT, in combination, or their
CC   complements. The methods are useful for diagnosing (a predisposition to)
CC   abnormally low levels of low high density lipoprotein-C (HDL-C) in a
CC   subject, The methods are useful in diagnosing (a predisposition to) or
CC   prognosticating diseases and disorders associated with abnormal lipid
CC   levels such as vascular and metabolic diseases, e.g., coronary artery
CC   disease. The present sequence is a human gene containing a SNP (single
CC   nucleotide polymorphism associated with low high density lipoprotein-C
CC   (HDL-C) levels.
XX
SQ   Sequence 160361 BP; 43435 A; 35277 C; 35459 G; 45990 T; 0 U; 200 Other;

  Query Match             34.1%;  Score 122.4;  DB 12;  Length 160361;
  Best Local Similarity   80.0%;  Pred. No. 3e-17;
  Matches  144;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||||| |||||||||||  || ||| ||||||||||||||||||||||
Db     157896 CCCGGCTAATTTTGTATTTTTTAGTAGAGGTGGGGTTTCTCCATGTTGGTCAGGCTGGTC 157955

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              || |||| |  |||||||||||||||||||| ||||||||||||| | |||||||||
Db     157956 TCAAACTCCCGACCTCAGGTGATCCGCCCGCCTTGGCCTCCCAAAGTGTTGGGATTACAG 158015

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||| |||||||||||| | ||   |  | |||  ||||  | | |     | ||   |||
Db     158016 GCGCGAGCCACCGCGCCCGGCTGAGCCAGACCACTTCTGTCTTTTAAGGCTCCTGCTAAT 158075


RESULT 9
AAK78202/c
ID   AAK78202 standard; DNA; 22651 BP.
XX
AC   AAK78202;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33014.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
```

PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 33014; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 22651 BP; 5939 A; 5107 C; 5360 G; 6245 T; 0 U; 0 Other;

Query Match 34.0%; Score 122.2; DB 4; Length 22651;
Best Local Similarity 90.9%; Pred. No. 2.3e-17;
Matches 130; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||||| |||||||||||| || ||| ||||||||||||||||||||||
Db       4383 CCCGGCTAATTTTGTATTTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTC 4324

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| | || |||||||||||||||||||| | |||||||||||| | |||||||
Db       4323 TCGAACTCCCAATCTCAGGTGATCCGCCCGCCTCAGTCTCCCAAAGTGCTGGAATTACAG 4264

Qy        121 GCGTGAGCCACCGCGCTCAGCCT 143
              ||||||||||||| | ||||||
Db       4263 GCGTGAGCCACCGCACCCAGCCT 4241
```

RESULT 10
ADZ70075/c
ID ADZ70075 standard; DNA; 313001 BP.
XX
AC ADZ70075;
XX
DT 14-JUL-2005 (first entry)
XX
DE Human insulin-like growth factor 1 receptor (IGF-IR) gene - SEQ ID 1.
XX
KW SNP detection; diagnosis; cardiac hypertrophy; cardiant;
KW insulin-like growth factor 1 receptor; gene; ds.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	5'UTR	3509. .3540
FT		/*tag= a
FT	CDS	3541. .311401
FT		/*tag= c
FT		/product= "Human IGF-IR protein"
FT	exon	3541. .3634
FT		/*tag= b
FT		/number= 1
FT	intron	3635. .61520
FT		/*tag= d
FT		/number= 1
FT	exon	61521. .62066
FT		/*tag= e
FT		/number= 2
FT	intron	62067. .245283
FT		/*tag= f
FT		/number= 2
FT	exon	245284. .245596
FT		/*tag= g
FT		/number= 3
FT	intron	245597. .250715
FT		/*tag= h
FT		/number= 3
FT	exon	250716. .250864
FT		/*tag= i
FT		/number= 4
FT	intron	250865. .253435
FT		/*tag= j
FT		/number= 4
FT	exon	253436. .253580
FT		/*tag= k
FT		/number= 5
FT	intron	253581. .262643
FT		/*tag= l
FT		/number= 5
FT	exon	262644. .262858

```
FT                   /*tag=  m
FT                   /number= 6
FT   intron          262859. .265273
FT                   /*tag=  n
FT                   /number= 6
FT   exon            265274. .265400
FT                   /*tag=  o
FT                   /number= 7
FT   intron          265401. .267002
FT                   /*tag=  p
FT                   /number= 7
FT   exon            267003. .267241
FT                   /*tag=  q
FT                   /number= 8
FT   intron          267242. .269922
FT                   /*tag=  r
FT                   /number= 8
FT   exon            269923. .270090
FT                   /*tag=  s
FT                   /number= 9
FT   intron          270091. .270630
FT                   /*tag=  t
FT                   /number= 9
FT   exon            270631. .270835
FT                   /*tag=  u
FT                   /number= 10
FT   intron          270836. .276106
FT                   /*tag=  v
FT                   /number= 10
FT   exon            276107. .276390
FT                   /*tag=  w
FT                   /number= 11
FT   intron          276391. .277834
FT                   /*tag=  x
FT                   /number= 11
FT   exon            277835. .277971
FT                   /*tag=  y
FT                   /number= 12
FT   intron          277972. .278483
FT                   /*tag=  z
FT                   /number= 12
FT   exon            278484. .278643
FT                   /*tag=  aa
FT                   /number= 13
FT   intron          278644. .283516
FT                   /*tag=  ab
FT                   /number= 13
FT   exon            283517. .283619
FT                   /*tag=  ac
FT                   /number= 14
FT   intron          283620. .284193
FT                   /*tag=  ad
FT                   /number= 14
FT   exon            284194. .284246
FT                   /*tag=  ae
FT                   /number= 15
FT   intron          284247. .288782
FT                   /*tag=  af
FT                   /number= 15
FT   exon            288783. .289012
FT                   /*tag=  ag
FT                   /number= 16
FT   intron          289013. .289274
FT                   /*tag=  ah
FT                   /number= 16
FT   exon            289275. .289385
FT                   /*tag=  ai
FT                   /number= 17
FT   intron          289386. .293159
FT                   /*tag=  aj
FT                   /number= 17
FT   exon            293160. .293319
FT                   /*tag=  ak
FT                   /number= 18
FT   intron          293320. .296881
FT                   /*tag=  al
FT                   /number= 18
FT   exon            296882. .297011
FT                   /*tag=  am
FT                   /number= 19
FT   intron          297012. .302532
FT                   /*tag=  an
FT                   /number= 19
FT   exon            302533. .302667
FT                   /*tag=  ao
FT                   /number= 20
FT   intron          302668. .311019
FT                   /*tag=  ap
FT                   /number= 20
FT   exon            311020. .311401
FT                   /*tag=  aq
FT                   /number= 21
FT   3'UTR           311402. .312241
FT                   /*tag=  ar
XX
PN   JP2005110607-A.
XX
PD   28-APR-2005.
XX
PF   09-OCT-2003; 2003JP-00350960.
XX
PR   09-OCT-2003; 2003JP-00350960.
XX
PA   (KOKU-) KOKURITSU JUNKANKI BYO CENT SOCHO.
PA   (DOKU-) DOKURITSU GYOSEI HOJIN IYAKUHIN IRYO KIK.
XX
DR   WPI; 2005-326229/34.
DR   P-PSDB; ADZ70089.
XX
PT   Testing hypertensive cardiac hypertrophy factor, by determining genotype
PT   of polymorphism in insulin-like growth factor 1 receptor (IGF-IR) gene of
PT   subject and estimating based on determined genotype.
XX
PS   Disclosure; SEQ ID NO 1; 19pp; Japanese.
XX
CC   The invention comprises a method of testing hypertensive hypertrophy
CC   factor. The method involves determining the genotype of a polymorphism in
CC   the insulin-like growth factor 1 receptor (IGF-IR) gene of a subject, and
CC   estimating the hypertensive cardiac hypertrophy factor based on the
CC   determined genotype. The method of the invention is useful for testing
CC   hypertensive cardiac hypertrophy factor, and thereby determining the risk
CC   of developing cardiac hypertrophy. The present DNA sequence represents
CC   the human IGF-IR gene.
XX
SQ   Sequence 313001 BP; 79741 A; 67002 C; 72630 G; 93626 T; 0 U; 2 Other;

  Query Match             34.0%;  Score 122;  DB 14;  Length 313001;
  Best Local Similarity   76.8%;  Pred. No. 4.2e-17;
  Matches  149;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              | || |||||||     |||||||||||| || ||| ||||||||||||||||||||| |
Db     165118 CTGGTTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTTT 165059

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              |||||| |  ||||||||||||||| |||||||  ||||| |||||||||  ||||||||||
Db     165058 CGAACTCCTGACCTCAGGTGATCCGGCCGCCTCAGCCTCTCAAAGTGCTGGGATTACAGG 164999

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAATG 181
              |||||||||||||  ||||| | |||    ||||||  ||||  |    ||      |
Db     164998 CGTGAGCCACCGCGTCCAGCCAGAGAAAGAATTTTCTAACATTCTAGGTTGTCTACAGAC 164939

Qy        182 CTTATGAAAACGAA 195
              ||||||| | ||
Db     164938 AATATGAAAGCAAA 164925
```

RESULT 11
ADS36450/c
ID ADS36450 standard; DNA; 321019 BP.
XX
AC ADS36450;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human autoimmune disease-related genomic DNA sequence - SEQ ID 1664.
XX
KW single nucleotide polymorphism detection; SNP detection;
KW rheumatoid arthritis; type 1 diabetes; multiple sclerosis;
KW systemic lupus erythematosus; inflammatory bowel disease; psoriasis;
KW thyroiditis; celiac disease; pernicious anaemia; asthma; vitiligo;
KW glomerulonephritis; Grave's disease; myocarditis; Sjogren's disease;
KW primary systemic vasculitis; ds.
XX
OS Homo sapiens.
XX
PN WO2004083403-A2.
XX
PD 30-SEP-2004.
XX
PF 18-MAR-2004; 2004WO-US008461.
XX
PR 18-MAR-2003; 2003US-0455444P.
PR 25-APR-2003; 2003US-0465241P.
XX
PA (APPL-) APPLERA CORP.
XX
PI Cargill M, Begovich AB, Alexander HC;
XX
DR WPI; 2004-728480/71.
XX
PT New isolated nucleic acid molecule comprises at least 8 contiguous
PT nucleotides where one of the nucleotides is a single nucleotide
PT polymorphism (SNP), useful for diagnosing or treating autoimmune
PT diseases, e.g. rheumatoid arthritis.
XX
PS Claim 16; SEQ ID NO 1664; 123pp; English.
XX
CC The invention comprises amino acid and coding sequences containing
CC genetic polymorphisms associated with an altered risk of developing an
CC autoimmune disease (e.g. rheumatoid arthritis). The invention further
CC comprises a method of identifying an individual that has an altered risk
CC of developing an autoimmune disease, comprising detecting a single
CC nucleotide polymorphism (SNP) in a nucleic acid of the invention. The DNA
CC and protein sequences of the invention are useful for diagnosing and
CC treating autoimmune diseases, such as: rheumatoid arthritis, type 1
CC diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory
CC bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious
CC anaemia, asthma, vitiligo, glomerulonephritis, Grave's disease,
CC myocarditis, Sjogren's disease, or primary systemic vasculitis. The
CC present nucleic acid represents a human autoimmune disease-related
CC genomic DNA sequence of the invention. NOTE: The present sequence is not
CC shown in the specification, but has been retrieved from the WIPO website.
XX
SQ Sequence 321019 BP; 81692 A; 68565 C; 74179 G; 95936 T; 0 U; 647 Other;

Query Match 34.0%; Score 122; DB 13; Length 321019;
Best Local Similarity 76.8%; Pred. No. 4.2e-17;
Matches 149; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

```
Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              | || |||||||     |||||||||||| || ||| ||||||||||||||||||| |
Db     167635 CTGGTTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTTT 167576

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              |||||| |  ||||||||||||||| ||||||| ||||| |||||||||| ||||||||||
Db     167575 CGAACTCCTGACCTCAGGTGATCCGGCCGCCTCAGCCTCTCAAAGTGCTGGGATTACAGG 167516

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAATG 181
              |||||||||||||||  ||||| | |||    |||||| |||| |    ||     |
Db     167515 CGTGAGCCACCGCGTCCAGCCAGAGAAAGAATTTTCTAACATTCTAGGTTGTCTACAGAC 167456

Qy        182 CTTATGAAAACGAA 195
              ||||||| | ||
Db     167455 AATATGAAAGCAAA 167442
```

RESULT 12
ABD32707/c
ID ABD32707 standard; DNA; 329019 BP.
XX
AC ABD32707;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human cancer-associated genomic DNA HD14-043.
XX
KW Human; ds; cancer-associated protein; gene; cytostatic; cancer;
KW leukaemia; lymphoma; CAP.
XX
OS Homo sapiens.
XX
PN WO2004074320-A2.
XX
PD 02-SEP-2004.
XX
PF 17-FEB-2004; 2004WO-US004730.
XX
PR 14-FEB-2003; 2003US-00367094.
PR 14-MAR-2003; 2003US-00388838.
PR 15-APR-2003; 2003US-00417375.
PR 13-JUN-2003; 2003US-00461862.
PR 15-SEP-2003; 2003US-00663431.
PR 15-DEC-2003; 2003US-00737318.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Morris DW, Malandro MS;
XX
DR WPI; 2004-652914/63.
XX
PT New isolated cancer-associated polynucleotides and polypeptides useful
PT for diagnosing, preventing or treating cancers, especially lymphoma and
PT leukemia, or in screening for agents that modulate cancer.
XX
PS claim 16; seqid 277; 310pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising at least 10
CC contiguous nucleotides of any of the 233 polynucleotide sequences given
CC in the specification, or its complement. The nucleic acids encode cancer-
CC associated proteins. Also included are an expression vector comprising
CC the isolated nucleic acid cited above, a host cell comprising the above
CC recombinant nucleic acid or expression vector, a microarray for detecting
CC a cancer-associated (CA) nucleic acid comprising at least one probe
CC comprising at least 10 contiguous nucleotides of any of the above-
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within
CC an open reading frame of a CA sequence selected from any of the 95
CC polynucleotide sequences as mentioned in the specification, or its
CC complement), an isolated antibody, (or its antigen binding fragment) that
CC binds to the above polypeptide, a hybridoma that produces the above
CC monoclonal antibody, a pharmaceutical composition comprising the above
CC antibody and a pharmaceutical excipient, a kit for detecting cancer
CC cells(comprising the antibody cited above, methods for diagnosing cancer
CC or for detecting the presence or absence of cancer cells in an
CC individual, a method for inhibiting growth of cancer cells in an
CC individual, a method for delivering a therapeutic agent to cancer cells
CC in an individual, an electronic library comprising the above
CC polynucleotide or polypeptide (or their fragments), methods of screening
CC for anticancer activity or for a bioactive agent capable of modulating
CC the activity of a CA protein (CAP), methods for detecting cancer
CC associated with expression of a polypeptide in a test cell sample, a
CC method for treating cancers and a method for inhibiting the expression of

```
CC   CA gene in a cell. The composition and methods are useful for detecting,
CC   diagnosing, preventing and treating cancers, especially lymphoma and
CC   leukaemia. These may also be used in screening for agents that modulate
CC   cancer. The present sequence is a human CAP genomic sequence. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 329019 BP; 84190 A; 70461 C; 76072 G; 98276 T; 0 U; 20 Other;

  Query Match            34.0%;  Score 122;  DB 13;  Length 329019;
  Best Local Similarity   76.8%;  Pred. No. 4.2e-17;
  Matches  149;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              | || |||||||     |||||||||||| || ||| |||||||||||||||||||| |
Db     171636 CTGGTTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTTT 171577

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              |||||| |  |||||||||||||||| ||||||| ||||| ||||||||| ||||||||||
Db     171576 CGAACTCCTGACCTCAGGTGATCCGGCCGCCTCAGCCTCTCAAAGTGCTGGGATTACAGG 171517

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAATG 181
              |||||||||||||  ||||| | |||    |||||| |||  |    ||     |
Db     171516 CGTGAGCCACCGCGTCCAGCCAGAGAAAGAATTTTCTAACATTCTAGGTTGTCTACAGAC 171457

Qy        182 CTTATGAAAACGAA 195
              ||||||| | ||
Db     171456 AATATGAAAGCAAA 171443


RESULT 13
ABL83966/c
ID   ABL83966 standard; cDNA; 381 BP.
XX
AC   ABL83966;
XX
DT   17-MAY-2002  (first entry)
XX
DE   Human ovarian cancer related cDNA clone SEQ ID NO:6944.
XX
KW   Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200192581-A2.
XX
PD   06-DEC-2001.
XX
PF   29-MAY-2001; 2001WO-US017756.
XX
PR   26-MAY-2000; 2000US-0207484P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Algate PA,  Harlocker SL,  Jones R;
XX
DR   WPI; 2002-122075/16.
XX
PT   Composition for therapy and diagnosis of ovarian cancer comprising
PT   polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT   polypeptide, antibody specific to polypeptide or T cell expressing
PT   polypeptide.
XX
PS   Claim 1; SEQ ID NO 6944; 489pp; English.
XX
CC   The present invention describes a composition (I) comprising: carriers
CC   and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC   polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC   from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC   (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC   or antigen presenting cells that express (II). (I) has cytostatic
CC   activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC   detecting ovarian cancer in a patient's biological sample preferably
CC   serum or ovarian tissue. The method comprises contacting a biological
CC   sample from a patient with (IV), detecting the amount of polynucleotide
CC   hybridising to (IV) and comparing the amount to a predetermined cutoff
CC   value and thereby detecting ovarian cancer in the patient, where the
CC   amount of polynucleotide hybridising to (IV) is detected preferably by
CC   polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC   useful for stimulating and/or expanding T cells specific for an ovarian
CC   tumour protein comprising contacting T cells with (III) or (II). (III) is
CC   useful in design and preparation of ribozyme molecules for inhibiting
CC   expression of the tumour polypeptides and proteins in tumour cells; and
CC   to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC   library using well known techniques
XX
SQ   Sequence 381 BP; 115 A; 83 C; 91 G; 92 T; 0 U; 0 Other;

  Query Match            33.9%;  Score 121.8;  DB 6;  Length 381;
  Best Local Similarity   77.8%;  Pred. No. 1.4e-17;
  Matches  147;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||     |||||||||||||||  ||| | ||||||||| ||||||||||
Db        328 CCCAGCTAATTTTTGTACTTTTAGTAGAGACGAGGTTTCGCCATGTTGGCCAGGCTGGTC 269

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| |||||||||||| |||||||||| |||||||||
Db        268 TCGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAG 209

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||||| ||   ||| |   |    ||||||| || ||| |   |    ||||
Db        208 GCGTGAGCCACCGTGCCTGGCCAGAATAGGTTTTTTCTTTCAACTTGATCAGTAGAAAAT 149

Qy        181 GCTTATGAA 189
              |   || ||
Db        148 GGACATCAA 140


RESULT 14
ACF91521/c
ID   ACF91521 standard; DNA; 546 BP.
XX
AC   ACF91521;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Human SIRS/sepsis diagnostic marker DNA fragment 10381.
XX
KW   Systemic inflammatory response syndrome; SIRS; antibacterial;
KW   immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004087949-A2.
XX
PD   14-OCT-2004.
XX
PF   31-MAR-2004; 2004WO-EP003419.
XX
PR   02-APR-2003; 2003DE-01015031.
PR   08-AUG-2003; 2003DE-01036511.
PR   02-SEP-2003; 2003DE-01040395.
XX
PA   (SIRS-) SIRS LAB GMBH.
XX
PI   Russwurm S,  Reinhart K,  Saluz H,  Straube E,  Zipfel PF,  Deigner H;
XX
DR   WPI; 2004-748070/73.
XX
PT   In vitro detection of systemic inflammatory response syndrome and related
PT   conditions, for e.g. monitoring progression, comprises detecting abnormal
PT   expression of disease-related genes.
XX
PS   Disclosure; Page; 75pp; German.
```

XX
CC The invention relates to a novel method for in vitro detection of
CC systemic inflammatory response syndrome (SIRS). The method comprises
CC detecting abnormal expression of disease-related genes, or their
CC associated peptides. The method of the invention demonstrates
CC antibacterial, immunosuppressive and antiinflammatory applications and
CC may be used for early differential diagnosis, monitoring progression,
CC assessing risk, assessing the likely response to treatment and for post
CC mortem diagnosis of systemic inflammatory response syndrome, sepsis and
CC sepsis-like conditions. The recombinant or synthetic nucleic acid
CC sequences of the invention, or derived proteins or peptides, may be
CC useful as calibrants in assays for the specified diseases, for evaluating
CC activity or toxicity in screening for active agents and/or for
CC preparation of agents for treatment or prevention of the specified
CC diseases. The current sequence is that of a human SIRS/sepsis diagnostic
CC marker DNA fragment of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at ftp.wipo.int/pub/published
CC pct_sequences. Furthermore, a number of arbitrary SEQ ID NO.s are
CC disclosed within the specification, however, these have not been taken
CC into account during indexing due to inconsistencies in application and
CC format
XX
SQ Sequence 546 BP; 180 A; 125 C; 122 G; 117 T; 0 U; 2 Other;

Query Match 33.9%; Score 121.8; DB 13; Length 546;
Best Local Similarity 77.8%; Pred. No. 1.4e-17;
Matches 147; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||    ||||||||||||||  ||| | |||||||||| ||||||||||
Db        329 CCCAGCTAATTTTTGTACTTTTAGTAGAGACGAGGTTTCGCCATGTTGGCCAGGCTGGTC 270

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| |||||||||||| |||||||||| ||||||||||
Db        269 TCGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAG 210

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||||| ||   ||| |   |    ||||||| || ||| |   |    ||||
Db        209 GCGTGAGCCACCGTGCCTGGCCAGAATAGGTTTTTTCTTTCAACTTGATCAGTAGAAAAT 150

Qy        181 GCTTATGAA 189
              |   || ||
Db        149 GGACATCAA 141
```

RESULT 15
AAS90976/c
ID AAS90976 standard; cDNA; 556 BP.
XX
AC AAS90976;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #26780.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;

XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG26789.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 26780; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 556 BP; 177 A; 129 C; 124 G; 126 T; 0 U; 0 Other;

Query Match 33.9%; Score 121.8; DB 5; Length 556;
Best Local Similarity 77.8%; Pred. No. 1.5e-17;
Matches 147; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||    ||||||||||||||  ||| | |||||||||| ||||||||||
Db        336 CCCAGCTAATTTTTGTACTTTTAGTAGAGACGAGGTTTCGCCATGTTGGCCAGGCTGGTC 277

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| |||||||||||| |||||||||| ||||||||||
Db        276 TCGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAG 217

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||||| ||   ||| |   |    ||||||| || ||| |   |    ||||
Db        216 GCGTGAGCCACCGTGCCTGGCCAGAATAGGTTTTTTCTTTCAACTTGATCAGTAGAAAAT 157

Qy        181 GCTTATGAA 189
              |   || ||
Db        156 GGACATCAA 148
```

Search completed: May 5, 2006, 17:45:50
Job time : 407 secs

GenCore version 5.1.8

OM nucleic - nucleic search, using sw model

Run on: May 6, 2006, 19:34:56 ; Search time 2648 Seconds
(without alignments)
6343.113 Million cell updates/sec

Title: US-10-009-579A-5_COPY_3188_3546
Perfect score: 359
Sequence: 1 cccggctaattttgtatctt.....ttttttatagtgttctggaa 359

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	124.4	34.7	1071	1	AL520459	AL520459 AL520459
	2	124.4	34.7	1686	4	CR619941	CR619941 full-leng
	3	124.4	34.7	1755	4	CR609780	CR609780 full-leng
c	4	123.8	34.5	2825	4	CR926093	CR926093 Pongo pyg
c	5	123.4	34.4	206	1	AA678616	AA678616 ah03c11.s
c	6	123	34.3	410	8	T96411	T96411 ye34e04.s1
	7	123	34.3	624	9	AQ419825	AQ419825 RPCI-11-1
c	8	122.4	34.1	202	1	AA368329	AA368329 EST79571
	9	121.8	33.9	350	7	CK823193	CK823193 ij25f02.y
c	10	121.8	33.9	354	1	AW078821	AW078821 xa99h09.x
c	11	121.8	33.9	355	1	AW511507	AW511507 xu57a01.x
	12	121.8	33.9	363	1	AW390284	AW390284 CM2-ST018
c	13	121.8	33.9	364	6	CB068575	CB068575 is30a03.x
c	14	121.8	33.9	366	1	AI347665	AI347665 qp01c06.x
c	15	121.8	33.9	376	1	AA644223	AA644223 ab63e10.s
c	16	121.8	33.9	381	1	AA581498	AA581498 nd18f02.s
c	17	121.8	33.9	393	1	AI472736	AI472736 ta13c04.x
c	18	121.8	33.9	398	2	BF475466	BF475466 nac30h01.
	19	121.8	33.9	411	3	BQ101225	BQ101225 ij25f02.y
c	20	121.8	33.9	421	1	AW440568	AW440568 xt15e04.x
	21	121.8	33.9	422	2	BE892611	BE892611 601433420
c	22	121.8	33.9	438	1	AA678932	AA678932 ah08g03.s
c	23	121.8	33.9	449	1	AW081610	AW081610 xc32b10.x
c	24	121.8	33.9	460	2	BE301610	BE301610 bb75e09.x
c	25	121.8	33.9	471	2	BE677244	BE677244 7d82d09.x
c	26	121.8	33.9	476	8	N67313	N67313 yz51e02.s1
c	27	121.8	33.9	493	5	BU198009	BU198009 EA25 Subt
	28	121.8	33.9	493	7	CN415426	CN415426 170004241
c	29	121.8	33.9	498	3	BM995211	BM995211 UI-H-ED0-
c	30	121.8	33.9	546	1	AI889995	AI889995 wm80c03.x
c	31	121.8	33.9	606	5	BQ778458	BQ778458 il31d07.x
	32	121.8	33.9	619	5	BX505458	BX505458 DKFZp686E
c	33	121.8	33.9	652	1	AA126814	AA126814 zn87c10.r
c	34	121.8	33.9	656	6	CA420015	CA420015 UI-H-FH0-
	35	121.8	33.9	712	6	CD246087	CD246087 AGENCOURT
c	36	121.8	33.9	721	6	CD364665	CD364665 UI-H-FT2-
	37	121.8	33.9	722	7	CN263776	CN263776 170004241
c	38	121.8	33.9	779	6	CA442904	CA442904 UI-H-DP0-
	39	121.8	33.9	3426	4	CR857101	CR857101 Pongo pyg
	40	121.8	33.9	3517	4	CR860781	CR860781 Pongo pyg
	41	121.6	33.9	660	3	BM997829	BM997829 UI-H-DI0-
	42	121.2	33.8	462	9	AQ559212	AQ559212 HS_2069_B
	43	121.2	33.8	511	2	BE464585	BE464585 hs85a05.x
c	44	120.8	33.6	419	7	CN264773	CN264773 170004245
c	45	120.8	33.6	468	2	BF931566	BF931566 IL2-NT020

ALIGNMENTS

RESULT 1
AL520459/c
LOCUS AL520459 1071 bp mRNA linear EST 24-MAR-2004
DEFINITION AL520459 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens cDNA clone CS0DB006YA12 3-PRIME, mRNA sequence.
ACCESSION AL520459
VERSION AL520459.3 GI:45695996
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1071)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:31038800.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6092.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DB006BA06NP1&c=6092.r.
FEATURES Location/Qualifiers
source 1. .1071
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB006YA12"
/tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match 34.7%; Score 124.4; DB 1; Length 1071;

```
  Best Local Similarity   86.7%;  Pred. No. 8.9e-13;
  Matches  137;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     ||||||||||||||| ||| ||||||||||||||||||||
Db        312 CCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCTCCATGTTGGTCAGGCTGGTC 253

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||| |||||||||||||||||||| |||||||||||| |||||||||
Db        252 TCGAACTCCTGACCTCATGTGATCCGCCCGCCTCGGCTTCCCAAAGTGCTGGGATTACAG 193

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCT 158
              |||||||| ||||||| ||||||| | | | | ||| |
Db        192 GCGTGAGCTACCGCGCCCAGCCTGCGTAGAGCCTTTTT 155


RESULT 2
CR619941
LOCUS       CR619941               1686 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DF033YI08 of Fetal brain of Homo sapiens
            (human).
ACCESSION   CR619941
VERSION     CR619941.1  GI:50500748
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1686)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1686)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1686
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF033YI08"
                     /tissue_type="Fetal brain"
                     /plasmid="pCMVSPORT_6"
ORIGIN

  Query Match            34.7%;  Score 124.4;  DB 4;  Length 1686;
  Best Local Similarity   86.7%;  Pred. No. 8.2e-13;
  Matches  137;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     ||||||||||||||| ||| ||||||||||||||||||||
Db       1404 CCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCTCCATGTTGGTCAGGCTGGTC 1463

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||| |||||||||||||||||||| |||||||||||| |||||||||
Db       1464 TCGAACTCCTGACCTCATGTGATCCGCCCGCCTCGGCTTCCCAAAGTGCTGGGATTACAG 1523

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCT 158
              |||||||| ||||||| ||||||| | | | | ||| |
Db       1524 GCGTGAGCTACCGCGCCCAGCCTGCGTAGAGCCTTTTT 1561


RESULT 3
CR609780
LOCUS       CR609780               1755 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DB006YA12 of Neuroblastoma Cot
            10-normalized of Homo sapiens (human).
ACCESSION   CR609780
VERSION     CR609780.1  GI:50490587
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1755)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1755)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1755
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DB006YA12"
                     /tissue_type="Neuroblastoma Cot 10-normalized"
                     /plasmid="pCMVSPORT_6"
ORIGIN

  Query Match            34.7%;  Score 124.4;  DB 4;  Length 1755;
  Best Local Similarity   86.7%;  Pred. No. 8.2e-13;
  Matches  137;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     ||||||||||||||| ||| ||||||||||||||||||||
Db       1444 CCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCTCCATGTTGGTCAGGCTGGTC 1503

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||| |||||||||||||||||||| |||||||||||| |||||||||
Db       1504 TCGAACTCCTGACCTCATGTGATCCGCCCGCCTCGGCTTCCCAAAGTGCTGGGATTACAG 1563

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCT 158
              |||||||| ||||||| ||||||| | | | | ||| |
Db       1564 GCGTGAGCTACCGCGCCCAGCCTGCGTAGAGCCTTTTT 1601


RESULT 4
CR926093/c
LOCUS       CR926093               2825 bp    mRNA    linear   HTC 06-DEC-2004
DEFINITION  Pongo pygmaeus mRNA; cDNA DKFZp459O2210 (from clone DKFZp459O2210).
ACCESSION   CR926093
VERSION     CR926093.1  GI:56403867
KEYWORDS    HTC.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 2825)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
```

Fobo,G., Han,M. and Wiemann,S.
CONSRTM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Agowa (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp459O2210) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp459O2210
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES Location/Qualifiers
source 1. .2825
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp459O2210"
/tissue_type="cortex"
/clone_lib="459 (synonym: pcor1). Vector pSport1_Sfi; host DH10B; sites SfilA + SfilB"
/dev_stage="adult"
/note="Gamma-tubulin complex component 4 (Homo sapiens), N-terminus truncated"
gene 1. .2825
/gene="DKFZp459O2210"
CDS <3. .1979
/gene="DKFZp459O2210"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAI29719.1"
/db_xref="GI:56403868"
/translation="LSGYPGSIFTWNKRSGLQVSQDFPFLHPSETSVLNRLCRLGTDY
IRFTEFIEQYTGHVQQQDHHPSQQGQGGLHGIYLRAFCTGLDSVLQPYRQALLDLEQE
FLGDPHLSISHVNYSLDQFQLLFPSVMVVVEQIKSQKIHGCQILETVYKHSCGGLPPV
RSALEKILAVCHGVMYKQLSAWMLHGLLLDQHEEFFIKQGPSSGNVSAQPEEDEEDLG
IGGLTGKQLRELQDLRLIEEENMLAPSLKQFSLRVEILPSYIPVRVAEKILFVGESVQ
MFENQNVNLTRKGSILKNQEDTFAAELHRLKQQPLFSLVDFEQVVDRIRSTVAEHLWK
LMVEESDLLGQLKIIKDFYLLGRGELFQAFIDTAQHMLKTPPTAVTEHDVNVAFQQSA
HKVLLDDDNLLPLLHLTIEYHGKEHKDGTQAREGPSRETSPREAPASGWAALGLSYKV
QWPLHILFTPAVLEKYNVVFKYLLSVRRVQAELLHCWALQMQRKHLKSNQTDAIKWRL
RNHMAFLVDNLQYYLQVDVLESQFSQLLHQINSTRDFESIRLAHDHFLSNLLAQSFIL
LKPVFHCLNEILDLCHSFCSLVSQNLGPLDERGAAQLSILVKGFSRQSSLLFKILSSV
RNHQINSDLAQLLLRLDYNKYYTQAGGTLGSFGM"

ORIGIN

Query Match 34.5%; Score 123.8; DB 4; Length 2825;
Best Local Similarity 69.9%; Pred. No. 9.6e-13;
Matches 167; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||  |||||||||      ||||||||||  ||| ||| |  |||||||||||||||||| |
Db       2659 CCTAGCTAATTTTTGTATTTTTAGTAGATTCGGGGTTTCATCATGTTGGTCAGGCTGGCC 2600

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              || ||||    || ||||||||||||||||||||||||||||||||||||| |||||||||
Db       2599 TCAAACTCTTGACTTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 2540

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              | ||||||||||  || | ||||  |||   ||||   || || ||||   |||||||| |
Db       2539 GTGTGAGCCACCATGCCCGGCCTAAGAAATACTTTAAGTATATTTTCATTAGCTAGAATT 2480

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTT 239
              ||  |        |   |  |||||  |  |  ||   | | |||||  ||   || | |
Db       2479 GCCCAATCTGTGTAGGTATAAATTACTTGGTATAGGGAGAGAGAAAGCCTATCTTACCT 2421
```

RESULT 5
AA678616/c
LOCUS AA678616 206 bp mRNA linear EST 02-DEC-1997

DEFINITION ah03c11.s1 Gessler Wilms tumor Homo sapiens cDNA clone IMAGE:1155572 3' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION AA678616
VERSION AA678616.1 GI:2659138
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 206)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 206.

FEATURES Location/Qualifiers
source 1. .206
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1155572"
/sex="pooled (6)"
/lab_host="DH10B"
/clone_lib="Gessler Wilms tumor"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; RNA was prepared from a pool of 6 anonymous Wilms' tumor RNAs. RNA was prepared by acid-phenol, followed by one round of oligo dT selection. cDNA library preparation was with the BRL/Life Tech. Superscript Plasmid system. An oligo-dT NotI primer for first strand synthesis generated gcggccgccc(t)n at the 3' end of the clones. A 5' SalI adaptor was used with sequence 5'-gtcgacccacgcgtccg-3'. Resulting cDNAs were size selected (average size 2 kb), NotI digested, and ligated into NotI/SalI-cut pSPORT1. Library was constructed by Dr. Manfred Gessler."

ORIGIN

Query Match 34.4%; Score 123.4; DB 1; Length 206;
Best Local Similarity 78.3%; Pred. No. 1.8e-12;
Matches 148; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||      ||||||||||||||||  ||| | |||||||||| ||||||||||
Db        192 CCCAGCTAATTTTTGTACTTTTAGTAGAGACGAGGTTTCGCCATGTTGGCCAGGCTGGTC 133

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| |||||||||||| |||||||||| |||||||||
Db        132 TCGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAG 73

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||||| ||   ||| |   |    ||||||| || ||| |  ||    ||||
Db         72 GCGTGAGCCACCGTGCCTGGCCAGTATAGGTTTTTTCTTTCAACTTGATCTGTAGAAAAT 13

Qy        181 GCTTATGAA 189
              |   || ||
Db         12 GGACATCAA 4
```

RESULT 6
T96411/c

```
LOCUS       T96411                  410 bp    mRNA    linear   EST 27-MAR-1995
DEFINITION  ye34e04.s1 Stratagene lung (#937210) Homo sapiens cDNA clone
            IMAGE:119646 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   T96411
VERSION     T96411.1  GI:735035
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 410)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 305
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -21m13
            High quality sequence stop: 305.
FEATURES             Location/Qualifiers
     source          1. .410
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:487935"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:119646"
                     /sex="male"
                     /dev_stage="72 years"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /clone_lib="Stratagene lung (#937210)"
                     /note="Organ: lung; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. normal lung.  Average insert size: 1.0 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             34.3%;  Score 123;  DB 8;  Length 410;
  Best Local Similarity   92.8%;  Pred. No. 1.9e-12;
  Matches  129;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCTCG 63
              ||||||||||||||| ||| |||||||||| | ||| ||||||||||||||||||||||||
Db        316 GGCTAATTTTGTATTTTTAAGTAGAGACAGGGTTTCTCCATGTTGGTCAGGCTGGTCTCG 257

Qy         64 AACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGGCG 123
              |||| | |||||||||||||||||| ||||||||||||||||||| ||||||||||||
Db        256 AACTCCCAACCTCAGGTGATCCGCCCACCTCGGCCTCCCAAAGTGCTCGGATTACAGGCG 197

Qy        124 TGAGCCACCGCGCTCAGCC 142
              ||||||||||||| |||||
Db        196 TGAGCCACCGCGCCCAGCC 178


RESULT 7
AQ419825
LOCUS       AQ419825                624 bp    DNA     linear   GSS 23-MAR-1999
DEFINITION  RPCI-11-179F14.TJ RPCI-11 Homo sapiens genomic clone
            RPCI-11-179F14, genomic survey sequence.
ACCESSION   AQ419825
VERSION     AQ419825.1  GI:4477549
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-179F14.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7568485"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-179F14"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             34.3%;  Score 123;  DB 9;  Length 624;
  Best Local Similarity   75.4%;  Pred. No. 1.8e-12;
  Matches  153;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              ||||||||||||     |||||||||||| || ||| | ||||||||| |||||| ||||
Db        257 CCGGCTAATTTTGTATTTTTAGTAGAGATGGAGTTTCGCCATGTTGGCCAGGCTAGTCT 316

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              ||||| |  |||||||||||||| || |||||||||||||||||||| ||||||||||
Db        317 TGAACTCCTGACCTCAGGTGATCCACCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGG 376

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAATG 181
              |||||||||| ||| |||||  |||| | ||||| |    |||     ||  | |
Db        377 CGTGAGCCACCACGCCCAGCCCAGGAAACTCCTTTCTAATGATTTCTTCCCCTCAATCTC 436

Qy        182 CTTATGAAAACGAAAAAAGAATT 204
              ||  |  ||| || ||||   ||
Db        437 CTGTTTGAAAGGAGAAAAAGTTT 459


RESULT 8
AA368329/c
LOCUS       AA368329                202 bp    mRNA    linear   EST 21-APR-1997
DEFINITION  EST79571 Placenta I Homo sapiens cDNA similar to EST containing Alu
            repeat, mRNA sequence.
ACCESSION   AA368329
```

```
VERSION     AA368329.1  GI:2020648
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 202)
  AUTHORS   Adams,M.D., Soares,M.B., Kerlavage,A.R., Fields,C. and Venter,J.C.
  TITLE     Rapid cDNA sequencing (expressed sequence tags) from a
            directionally cloned human infant brain cDNA library
  JOURNAL   Nat. Genet. 4, 373-380 (1993)
   PUBMED   8401585
COMMENT     Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .202
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="ATCC (inhost):173016"
                     /db_xref="taxon:9606"
                     /tissue_type="placenta"
                     /dev_stage="fetus"
                     /clone_lib="Placenta I"
                     /note="Organ: placenta; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: EcoRI"
ORIGIN

  Query Match             34.1%;  Score 122.4;  DB 1;  Length 202;
  Best Local Similarity   86.5%;  Pred. No. 2.8e-12;
  Matches  135;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     |||||||||||||| || ||| ||||||||||||||||||||||
Db        170 CCCGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTC 111

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| |||| ||||||||||||||||||| |||||||||
Db        110 TCGAACTCCTGACCTCAGGTGATCTGCCCACCTCGGCCTCCCAAAGTGCTGGGATTACAG 51

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTT 156
              || ||||||||||||||   ||||| || || |||||
Db         50 GCATGAGCCACCGCGCCTGGCCTGAGACCATCTTTT 15


RESULT 9
CK823193
LOCUS       CK823193                350 bp    mRNA    linear   EST 11-MAR-2004
DEFINITION  ij25f02.y5 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
            cDNA clone IMAGE:6135699 5', mRNA sequence.
ACCESSION   CK823193
VERSION     CK823193.1  GI:44840118
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 350)
  AUTHORS   Melton,D., Meadows,A., Clifton,S., Hillier,L., Marra,M., Pape,D.,
            Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B.,
            Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M.,
            McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and
            Bowers,Y.
  TITLE     WashU-Harvard Pancreas EST Project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            This read is a 5' RESEQUENCE of a previously sequenced pancreas
            clone
            This read has been verified (found to hit its original self in the
            correct orientation)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .350
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6135699"
                     /sex="Both"
                     /tissue_type="Islets of Langerhans"
                     /dev_stage="Adult"
                     /lab_host="DH10B"
                     /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
                     /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
                     Site_2: Sal 1; Starting library constructed using
                     SuperScript Plasmid Library kit (Life Technologies). cDNA
                     made by oligo-dT priming. Size-selected by column
                     fractionation; average insert size 1.08 kb. Library was
                     amplified once on solid support and plasmid DNA from
                     library was prepared. The library DNA was normalized by
                     method #4 from Bonaldo, Lennon, and Soares 1996 Genome
                     Research 6:791-806; 0.5 microgram single-stranded library
                     plasmid DNA was mixed with 5 micrograms PCR product
                     representing library inserts and hybridized to an Ecot of
                     20. Single-stranded (unhybridized) plasmids were isolated
                     by hydroxyapatite chromatography and used to make this
                     library."
ORIGIN

  Query Match             33.9%;  Score 121.8;  DB 7;  Length 350;
  Best Local Similarity   77.8%;  Pred. No. 3.2e-12;
  Matches  147;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||     |||||||||||||||  ||| | ||||||||| ||||||||||
Db         58 CCCAGCTAATTTTTGTACTTTTAGTAGAGACGAGGTTTCGCCATGTTGGCCAGGCTGGTC 117

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| |||||||||||| |||||||||| |||||||||
Db        118 TCGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAG 177

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||||| ||   ||| |   |    ||||||| || ||| |   |    ||||
Db        178 GCGTGAGCCACCGTGCCTGGCCAGAATAGGTTTTTTCTTTCAACTTGATCAGTAGAAAAT 237

Qy        181 GCTTATGAA 189
              |   || ||
Db        238 GGACATCAA 246


RESULT 10
AW078821/c
LOCUS       AW078821                354 bp    mRNA    linear   EST 14-OCT-1999
DEFINITION  xa99h09.x1 NCI_CGAP_Co17 Homo sapiens cDNA clone IMAGE:2574977 3'
            similar to contains Alu repetitive element;contains element MER20
            repetitive element ;, mRNA sequence.
ACCESSION   AW078821
VERSION     AW078821.1  GI:6033973
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 354)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D.,Ph.D. cDNA Library Preparation: Life
            Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
            I.M.A.G.E. Consortium DNA Sequencing by: Washington University
            Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 341.
FEATURES             Location/Qualifiers
     source          1. .354
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2574977"
                     /tissue_type="juvenile granulosa tumor"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Co17"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies."
ORIGIN

  Query Match             33.9%;  Score 121.8;  DB 1;  Length 354;
  Best Local Similarity   77.8%;  Pred. No. 3.2e-12;
  Matches  147;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||    |||||||||||||||  ||| | ||||||||| ||||||||||
Db        328 CCCAGCTAATTTTTGTACTTTTAGTAGAGACGAGGTTTCGCCATGTTGGCCAGGCTGGTC 269

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| |||||||||||| |||||||||| |||||||||
Db        268 TCGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAG 209

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||||||||| ||   ||| |   |    ||||||| || ||| |   |    ||||
Db        208 GCGTGAGCCACCGTGCCTGGCCAGAATAGGTTTTTTCTTTCAACTTGATCAGTAGAAAAT 149

Qy        181 GCTTATGAA 189
              |   || ||
Db        148 GGACATCAA 140


RESULT 11
AW511507/c
LOCUS       AW511507                355 bp    mRNA    linear   EST 03-MAR-2000
DEFINITION  xu57a01.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2805768 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AW511507
VERSION     AW511507.1  GI:7149509
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 355)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .355
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2805768"
                     /tissue_type="well-differentiated endometrial
                     adenocarcinoma, 7 pooled tumors"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Ut1"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb. Life Technologies catalog #:
                     11538-014"
ORIGIN

  Query Match             33.9%;  Score 121.8;  DB 1;  Length 355;
  Best Local Similarity   77.8%;  Pred. No. 3.2e-12;
  Matches  147;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||    |||||||||||||||  ||| | ||||||||| ||||||||||
Db        328 CCCAGCTAATTTTTGTACTTTTAGTAGAGACGAGGTTTCGCCATGTTGGCCAGGCTGGTC 269

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| |||||||||||| |||||||||| |||||||||
Db        268 TCGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAG 209

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||||||||| ||   ||| |   |    ||||||| || ||| |   |    ||||
Db        208 GCGTGAGCCACCGTGCCTGGCCAGAATAGGTTTTTTCTTTCAACTTGATCAGTAGAAAAT 149

Qy        181 GCTTATGAA 189
              |   || ||
Db        148 GGACATCAA 140


RESULT 12
AW390284
LOCUS       AW390284                363 bp    mRNA    linear   EST 04-FEB-2000
DEFINITION  CM2-ST0182-221099-023-f05 ST0182 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW390284
VERSION     AW390284.1  GI:6894943
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 363)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
```

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-ST0182-221099-023-f05&t3=1999-10-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 363.

FEATURES Location/Qualifiers
source 1. .363
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="ST0182"
/note="Organ: stomach; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 33.9%; Score 121.8; DB 1; Length 363;
Best Local Similarity 77.8%; Pred. No. 3.2e-12;
Matches 147; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||    ||||||||||||||||  ||| | |||||||||| ||||||||||
Db         62 CCCAGCTAATTTTTGTACTTTTAGTAGAGACGAGGTTTCGCCATGTTGGCCAGGCTGGTC 121

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| |||||||||||| |||||||||| |||||||||
Db        122 TCGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAG 181

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||||||||| ||   ||| |   |     ||||||| || ||| |   |    ||||
Db        182 GCGTGAGCCACCGTGCCTGGCCAGAATAGGTTTTTTCTTTCAACTTGATCAGTAGAAAAT 241

Qy        181 GCTTATGAA 189
              |   || ||
Db        242 GGACATCAA 250
```

RESULT 13
CB068575/c
LOCUS CB068575 364 bp mRNA linear EST 21-JAN-2003
DEFINITION is30a03.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553806 3', mRNA sequence.
ACCESSION CB068575
VERSION CB068575.1 GI:27813095
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 364)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 348.

FEATURES Location/Qualifiers
source 1. .364
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6553806"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match 33.9%; Score 121.8; DB 6; Length 364;
Best Local Similarity 77.8%; Pred. No. 3.2e-12;
Matches 147; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||    ||||||||||||||||  ||| | |||||||||| ||||||||||
Db        326 CCCAGCTAATTTTTGTACTTTTAGTAGAGACGAGGTTTCGCCATGTTGGCCAGGCTGGTC 267

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| |||||||||||| |||||||||| |||||||||
Db        266 TCGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAG 207

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||||||||| ||   ||| |   |     ||||||| || ||| |   |    ||||
Db        206 GCGTGAGCCACCGTGCCTGGCCAGAATAGGTTTTTTCTTTCAACTTGATCAGTAGAAAAT 147

Qy        181 GCTTATGAA 189
              |   || ||
Db        146 GGACATCAA 138
```

RESULT 14
AI347665/c
LOCUS AI347665 366 bp mRNA linear EST 02-FEB-1999
DEFINITION qp01c06.x1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1916746 3' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION AI347665
VERSION AI347665.1 GI:4084871
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 366)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 436 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES Location/Qualifiers
source 1. .366
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1916746"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid5"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCAATATTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Patima Bonaldo. "

ORIGIN

Query Match 33.9%; Score 121.8; DB 1; Length 366;
Best Local Similarity 77.8%; Pred. No. 3.2e-12;
Matches 147; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||    ||||||||||||||  ||| | ||||||||| ||||||||||
Db        336 CCCAGCTAATTTTTGTACTTTTAGTAGAGACGAGGTTTCGCCATGTTGGCCAGGCTGGTC 277

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| |||||||||||| |||||||||| |||||||||
Db        276 TCGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAG 217

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||||| ||   ||| |   |    ||||||| || ||| |   |    ||||
Db        216 GCGTGAGCCACCGTGCCTGGCCAGAATAGGTTTTTTCTTTCAACTTGATCAGTAGAAAAT 157

Qy        181 GCTTATGAA 189
              |   || ||
Db        156 GGACATCAA 148
```

RESULT 15
AA644223/c
LOCUS AA644223 376 bp mRNA linear EST 06-MAR-1998
DEFINITION ab63e10.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone IMAGE:845514 3' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION AA644223
VERSION AA644223.1 GI:2569441
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 376)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 534 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.

FEATURES Location/Qualifiers
source 1. .376
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:845514"
/tissue_type="lung carcinoma"
/cell_line="NCI-H69"
/dev_stage="cell line NCI-H69"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene lung carcinoma 937218"
/note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Small cell carcinoma cell line NCI-H69. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"

ORIGIN

Query Match 33.9%; Score 121.8; DB 1; Length 376;
Best Local Similarity 77.8%; Pred. No. 3.2e-12;
Matches 147; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||    ||||||||||||||  ||| | ||||||||| ||||||||||
Db        330 CCCAGCTAATTTTTGTACTTTTAGTAGAGACGAGGTTTCGCCATGTTGGCCAGGCTGGTC 271

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| |  |||||||||||||| |||||||||||| |||||||||| |||||||||
Db        270 TCGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAG 211

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||||| ||   ||| |   |    ||||||| || ||| |   |    ||||
Db        210 GCGTGAGCCACCGTGCCTGGCCAGAATAGGTTTTTTCTTTCAACTTGATCAGTAGAAAAT 151

Qy        181 GCTTATGAA 189
              |   || ||
Db        150 GGACATCAA 142
```

Search completed: May 6, 2006, 20:53:42
Job time : 2657 secs

GenCore version 5.1.8

OM nucleic - nucleic search, using sw model

Run on: May 6, 2006, 10:51:09 ; Search time 137 Seconds
(without alignments)
4657.990 Million cell updates/sec

Title: US-10-009-579A-5_COPY_3188_3546
Perfect score: 359
Sequence: 1 cccggctaattttgtatctt.....ttttttatagtgttctggaa 359

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	123.8	34.5	38346	3	US-09-949-016-14502	Sequence 14502, A
	2	123.2	34.3	360470	3	US-09-949-016-13173	Sequence 13173, A
c	3	122.8	34.2	41182	3	US-09-949-016-17457	Sequence 17457, A
c	4	122.6	34.2	47741	3	US-09-949-002-785	Sequence 785, App
c	5	122.2	34.0	31618	3	US-09-949-016-14759	Sequence 14759, A
c	6	122	34.0	601	3	US-09-949-016-25488	Sequence 25488, A
c	7	122	34.0	601	3	US-09-949-016-25489	Sequence 25489, A
c	8	122	34.0	601	3	US-09-949-016-25490	Sequence 25490, A
c	9	122	34.0	601	3	US-09-949-016-25491	Sequence 25491, A
c	10	122	34.0	601	3	US-09-949-016-73666	Sequence 73666, A
c	11	122	34.0	601	3	US-09-949-016-73667	Sequence 73667, A
c	12	122	34.0	601	3	US-09-949-016-73668	Sequence 73668, A
c	13	122	34.0	601	3	US-09-949-016-73669	Sequence 73669, A
c	14	122	34.0	39489	3	US-09-949-016-13886	Sequence 13886, A
c	15	122	34.0	70828	3	US-09-949-016-12122	Sequence 12122, A
	16	121.8	33.9	601	3	US-09-949-016-181612	Sequence 181612,
c	17	121.8	33.9	69709	3	US-09-949-016-15784	Sequence 15784, A
	18	121.8	33.9	77994	3	US-09-949-016-12517	Sequence 12517, A
	19	121.8	33.9	77994	3	US-09-949-016-16021	Sequence 16021, A
	20	121.8	33.9	125672	3	US-09-949-016-16956	Sequence 16956, A
c	21	121.8	33.9	146039	3	US-09-949-016-12449	Sequence 12449, A
c	22	121.4	33.8	601	3	US-09-949-016-178724	Sequence 178724,
	23	121.4	33.8	35688	3	US-09-949-016-16873	Sequence 16873, A
c	24	121.2	33.8	601	3	US-09-949-016-186521	Sequence 186521,
	25	121.2	33.8	25260	3	US-09-949-016-11985	Sequence 11985, A
	26	121.2	33.8	25260	3	US-09-949-016-12907	Sequence 12907, A
	27	121.2	33.8	60589	3	US-09-949-016-17070	Sequence 17070, A
c	28	120.8	33.6	20662	3	US-09-949-016-16717	Sequence 16717, A
	29	120.6	33.6	23521	3	US-09-949-016-14592	Sequence 14592, A
c	30	120.6	33.6	133278	3	US-09-949-016-12524	Sequence 12524, A
	31	120.4	33.5	60137	3	US-09-949-016-14735	Sequence 14735, A
	32	120.4	33.5	60137	3	US-09-949-016-14912	Sequence 14912, A
	33	120.4	33.5	120727	3	US-09-949-016-15787	Sequence 15787, A
	34	120.4	33.5	120727	3	US-09-949-016-15788	Sequence 15788, A
	35	120.2	33.5	102409	3	US-09-949-016-15148	Sequence 15148, A
c	36	120.2	33.5	134434	3	US-09-949-016-17362	Sequence 17362, A
c	37	119.8	33.4	42741	3	US-09-949-016-11857	Sequence 11857, A
c	38	119.8	33.4	45819	3	US-09-949-002-825	Sequence 825, App
c	39	119.6	33.3	41941	3	US-09-949-016-17380	Sequence 17380, A
c	40	119.6	33.3	42975	3	US-09-949-016-11965	Sequence 11965, A
	41	119.4	33.3	44554	3	US-09-949-016-12787	Sequence 12787, A
	42	119.4	33.3	44555	3	US-09-949-016-12043	Sequence 12043, A
c	43	119.2	33.2	601	3	US-09-949-016-162151	Sequence 162151,
c	44	119.2	33.2	8220	3	US-09-797-908-3	Sequence 3, Appli
	45	119.2	33.2	15084	3	US-09-949-016-16277	Sequence 16277, A

ALIGNMENTS

RESULT 1
US-09-949-016-14502/c
; Sequence 14502, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14502
; LENGTH: 38346
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(38346)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14502

Query Match 34.5%; Score 123.8; DB 3; Length 38346;
Best Local Similarity 83.8%; Pred. No. 1.6e-21;
Matches 140; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||     |||||||||||| || ||| ||||||||||||||||||||
Db      21255 CCCAGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTC 21196

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| || |||||||||||||||||||||||| | |||||||||||| |||||||||
Db      21195 TCGAACTCCAGACCTCAGGTGATCCGCCCGCCTCAGTCTCCCAAAGTGCTGGGATTACAG 21136

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTC 167
              |||||||||| |||| | ||||| || || ||||   |||  ||
Db      21135 GCGTGAGCCACTGCGCCCGGCCTGATAAAACTTTTTAAAACAGGCTC 21089
```

RESULT 2

```
US-09-949-016-13173
; Sequence 13173, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13173
;   LENGTH: 360470
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13173

  Query Match             34.3%;  Score 123.2;  DB 3;  Length 360470;
  Best Local Similarity   69.6%;  Pred. No. 4.7e-21;
  Matches  167;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              | ||||||||||      ||||||||||| || ||| | ||||||||| ||| |||||||
Db     187435 CTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGCCTGGTCT 187494

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              ||||||  |  |||||||||||  |||||||||||||||||||||||||| |||||||||
Db     187495 CGAACTCCTGACCTCAGGTGATTTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGA 187554

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAATG 181
              ||||||||||  || |||||  |     | |||   ||  | ||  | || |     ||
Db     187555 CGTGAGCCACCATGCCCAGCCGTGTTTTTCTTTTAGATAACTTTTGTATTGTTTTTCATA 187614

Qy        182 CTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTCC 241
              | |      || ||| | |||      ||||| || |  |||  | |  |||  | || ||
Db     187615 CATTCAGTAAAGAATATAGATAAGCAAAGAGAAAATTAAAATTACCCTTCAGCTCCTCCC 187674


RESULT 3
US-09-949-016-17457/c
; Sequence 17457, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17457
;   LENGTH: 41182
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(41182)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17457

  Query Match             34.2%;  Score 122.8;  DB 3;  Length 41182;
  Best Local Similarity   74.8%;  Pred. No. 3e-21;
  Matches  154;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||||||| |||||||||||||||| ||| |||||||||||||||| |||||
Db      16778 CCCCGCTAATTTTGTATTTTTTAGTAGAGACGGGGTTTCTCCATGTTGGTCAGGTTGGTC 16719

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              || |||| | |||||||||||||| ||||||||| |||||||||||||||| |||||||||
Db      16718 TCAAACTCCCAACCTCAGGTGATCTGCCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAG 16659

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||||||| | |||  | ||   || |  |  |  |  |  |  |  |
Db      16658 GCGTGAGCCACCGCGCCCGGCCAAGAAAATTTTTCTTCAAAGTGAATATATTGCATTACA 16599

Qy        181 GCTTATGAAAACGAAAAAAGAATTAT 206
              ||    || | || |||  |   ||
Db      16598 CCTGTCAAAGAAGACAAATAACCAAT 16573


RESULT 4
US-09-949-002-785/c
; Sequence 785, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785
;   LENGTH: 47741
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(47741)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-002-785

  Query Match             34.2%;  Score 122.6;  DB 3;  Length 47741;
  Best Local Similarity   70.4%;  Pred. No. 3.5e-21;
  Matches  164;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||      |||||||||||||||| ||| | ||||||||| ||||| ||||
Db      10107 CCCGGCTAATTTTTGTGGTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCCGGTC 10048

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              || |||| |  |||||||||||||||| |||||||||||||||| ||||| ||||||| |
Db      10047 TCAAACTCCTGACCTCAGGTGATCCGCCTGCCTCGGCCTCCCAACGTGCTGGGATTACGG 9988

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              | ||||||||| |||| |||||||  | | | ||||  |   | |  ||||  || || |
Db       9987 GTGTGAGCCACTGCGCCCAGCCTGATACCTCATTTTTCTTACTATGAAAGTAATACAATT 9928

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCAT 233
              |  |    | ||| | |   |  |  ||  |||   |||    |||
Db       9927 CGAAAACATGGAGGAAACAAATGGAAAACTATAAAGAATATTAAAAATTACAT 9875


RESULT 5
US-09-949-016-14759/c
; Sequence 14759, Application US/09949016
; Patent No. 6812339
```

```
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14759
;   LENGTH: 31618
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14759

  Query Match             34.0%;  Score 122.2;  DB 3;  Length 31618;
  Best Local Similarity   83.8%;  Pred. No. 3.9e-21;
  Matches  150;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||||| |||||||||||| || ||| |||||| ||||||||| |||||
Db       9317 CCCGGCTAATTTTGTAT-TTTTAGTAGAGATGGGGTTTCTCCATATTGGTCAGGTTGGTC 9259

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| | |||||||||||||||| |||||||||||||||||||||||||| |||||||||
Db       9258 TCGAACTCCTAACCTCAGGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 9199

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAA 179
              |||||||||||||| || |||||||| | | || | || |||    |     | ||  || ||
Db       9198 GCGTGAGCCACCGAGCCCAGCCTGAGGAAACATGTTTTTAAGCCACAGAATGACAGTAA 9140


RESULT 6
US-09-949-016-25488/c
; Sequence 25488, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25488
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-25488

  Query Match             34.0%;  Score 122;  DB 3;  Length 601;
  Best Local Similarity   84.6%;  Pred. No. 1.3e-21;
  Matches  137;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||     |||||||||||||| ||| | ||||||||||||||||||||
Db        484 CCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTC 425

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| |  |||||||||||||| ||||||||||||||||||||||||| |||||||||
Db        424 TTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 365

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACA 162
              |||||||||||||||| | ||| |      || |||||| | |
Db        364 GCGTGAGCCACCGCGCCCGGCCAGTTTTTACTTTTTCTAAAA 323


RESULT 7
US-09-949-016-25489/c
; Sequence 25489, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25489
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-25489

  Query Match             34.0%;  Score 122;  DB 3;  Length 601;
  Best Local Similarity   84.6%;  Pred. No. 1.3e-21;
  Matches  137;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||     |||||||||||||| ||| | ||||||||||||||||||||
Db        318 CCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTC 259

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| |  |||||||||||||| ||||||||||||||||||||||||| |||||||||
Db        258 TTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 199

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACA 162
              |||||||||||||||| | ||| |      || |||||| | |
Db        198 GCGTGAGCCACCGCGCCCGGCCAGTTTTTACTTTTTCTAAAA 157


RESULT 8
US-09-949-016-25490/c
; Sequence 25490, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25490
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-25490
```

```
  Query Match             34.0%;  Score 122;  DB 3;  Length 601;
  Best Local Similarity   84.6%;  Pred. No. 1.3e-21;
  Matches  137;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     |||||||||||||||| ||| | ||||||||||||||||||||
Db        315 CCCGGCTAATTTTTRTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTC 256

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| |  |||||||||||||| ||||||||||||||||||||||||| |||||||||
Db        255 TTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 196

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACA 162
              |||||||||||||||| | ||| |      || |||||| | |
Db        195 GCGTGAGCCACCGCGCCCGGCCAGTTTTTACTTTTTCTAAAA 154


RESULT 9
US-09-949-016-25491/c
; Sequence 25491, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25491
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-25491

  Query Match             34.0%;  Score 122;  DB 3;  Length 601;
  Best Local Similarity   84.6%;  Pred. No. 1.3e-21;
  Matches  137;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     |||||||||||||||| ||| | ||||||||||||||||||||
Db        165 CCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTC 106

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| |  |||||||||||||| ||||||||||||||||||||||||| |||||||||
Db        105 TTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 46

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACA 162
              |||||||||||||||| | ||| |      || |||||| | |
Db         45 GCGTGAGCCACCGCGCCCGGCCAGTTTTTACTTTTTCTAAAA 4


RESULT 10
US-09-949-016-73666/c
; Sequence 73666, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73666
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-73666

  Query Match             34.0%;  Score 122;  DB 3;  Length 601;
  Best Local Similarity   84.6%;  Pred. No. 1.3e-21;
  Matches  137;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     |||||||||||||||| ||| | ||||||||||||||||||||
Db        484 CCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTC 425

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| |  |||||||||||||| ||||||||||||||||||||||||| |||||||||
Db        424 TTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 365

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACA 162
              |||||||||||||||| | ||| |      || |||||| | |
Db        364 GCGTGAGCCACCGCGCCCGGCCAGTTTTTACTTTTTCTAAAA 323


RESULT 11
US-09-949-016-73667/c
; Sequence 73667, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73667
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-73667

  Query Match             34.0%;  Score 122;  DB 3;  Length 601;
  Best Local Similarity   84.6%;  Pred. No. 1.3e-21;
  Matches  137;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     |||||||||||||||| ||| | ||||||||||||||||||||
Db        318 CCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTC 259

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| |  |||||||||||||| ||||||||||||||||||||||||| |||||||||
Db        258 TTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 199

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACA 162
              |||||||||||||||| | ||| |      || |||||| | |
Db        198 GCGTGAGCCACCGCGCCCGGCCAGTTTTTACTTTTTCTAAAA 157


RESULT 12
US-09-949-016-73668/c
```

```
; Sequence 73668, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73668
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-73668

  Query Match             34.0%;  Score 122;  DB 3;  Length 601;
  Best Local Similarity   84.6%;  Pred. No. 1.3e-21;
  Matches  137;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     |||||||||||||||| ||| | ||||||||||||||||||||
Db        315 CCCGGCTAATTTTTRTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTC 256

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| |  |||||||||||||| ||||||||||||||||||||||||| |||||||||
Db        255 TTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 196

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACA 162
              |||||||||||||||| | ||| |      || |||||| | |
Db        195 GCGTGAGCCACCGCGCCCGGCCAGTTTTTACTTTTTCTAAAA 154


RESULT 13
US-09-949-016-73669/c
; Sequence 73669, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73669
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-73669

  Query Match             34.0%;  Score 122;  DB 3;  Length 601;
  Best Local Similarity   84.6%;  Pred. No. 1.3e-21;
  Matches  137;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     |||||||||||||||| ||| | ||||||||||||||||||||
Db        165 CCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTC 106

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| |  |||||||||||||| ||||||||||||||||||||||||| |||||||||
Db        105 TTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 46

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACA 162
              |||||||||||||||| | ||| |      || |||||| | |
Db         45 GCGTGAGCCACCGCGCCCGGCCAGTTTTTACTTTTTCTAAAA 4


RESULT 14
US-09-949-016-13886/c
; Sequence 13886, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13886
;   LENGTH: 39489
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(39489)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13886

  Query Match             34.0%;  Score 122;  DB 3;  Length 39489;
  Best Local Similarity   84.6%;  Pred. No. 4.7e-21;
  Matches  137;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     |||||||||||||||| ||| | ||||||||||||||||||||
Db      20196 CCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTC 20137

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| |  |||||||||||||| ||||||||||||||||||||||||| |||||||||
Db      20136 TTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 20077

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACA 162
              |||||||||||||||| | ||| |      || |||||| | |
Db      20076 GCGTGAGCCACCGCGCCCGGCCAGTTTTTACTTTTTCTAAAA 20035


RESULT 15
US-09-949-016-12122/c
; Sequence 12122, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12122
;   LENGTH: 70828
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(70828)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12122

  Query Match             34.0%;  Score 122;  DB 3;  Length 70828;
  Best Local Similarity   84.6%;  Pred. No. 5.7e-21;
  Matches  137;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||     |||||||||||||||| ||| | ||||||||||||||||||||
Db       3535 CCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTC 3476

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| |  |||||||||||||| ||||||||||||||||||||||||| |||||||||
Db       3475 TTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 3416

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACA 162
              ||||||||||||||| | ||| |     || |||||| | |
Db       3415 GCGTGAGCCACCGCGCCCGGCCAGTTTTTACTTTTTCTAAAA 3374


Search completed: May  6, 2006, 10:53:37
Job time : 141 secs
```

GenCore version 5.1.8

OM nucleic - nucleic search, using sw model

Run on: May 6, 2006, 12:06:01 ; Search time 575 Seconds
(without alignments)
5162.968 Million cell updates/sec

Title: US-10-009-579A-5_COPY_3188_3546
Perfect score: 359
Sequence: 1 cccggctaatttttgtatctt.....ttttttatagtgttctggaa 359

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	359	100.0	4282	5	US-10-009-579-5	Sequence 5, Appli
c	2	124	34.5	2041	4	US-09-925-065A-72033	Sequence 72033, A
c	3	124	34.5	2041	4	US-09-925-065A-72034	Sequence 72034, A
c	4	124	34.5	2041	4	US-09-925-065A-72035	Sequence 72035, A
	5	123.8	34.5	438	8	US-10-357-930-13830	Sequence 13830, A
c	6	123.6	34.4	1136	5	US-10-027-632-117277	Sequence 117277,
c	7	123.6	34.4	1136	5	US-10-027-632-117278	Sequence 117278,
c	8	123.6	34.4	1136	6	US-10-027-632-117277	Sequence 117277,
c	9	123.6	34.4	1136	6	US-10-027-632-117278	Sequence 117278,
c	10	123	34.3	53623	8	US-10-417-375-44	Sequence 44, Appl
	11	122.4	34.1	160361	7	US-10-235-192A-35	Sequence 35, Appl
	12	122	34.0	464	4	US-09-925-065A-591033	Sequence 591033,
	13	122	34.0	1031	4	US-09-925-065A-724096	Sequence 724096,
c	14	122	34.0	321019	8	US-10-741-600-17566	Sequence 17566, A
c	15	122	34.0	329019	8	US-10-388-838-48	Sequence 48, Appl
c	16	121.8	33.9	381	3	US-09-867-701-6944	Sequence 6944, Ap
	17	121.8	33.9	424	4	US-09-925-065A-47240	Sequence 47240, A
c	18	121.8	33.9	493	9	US-10-933-118-33	Sequence 33, Appl
c	19	121.8	33.9	556	9	US-10-450-763-26780	Sequence 26780, A
	20	121.8	33.9	3470	8	US-10-357-930-25055	Sequence 25055, A
	21	121.8	33.9	122673	9	US-10-737-082-33	Sequence 33, Appl
	22	121.8	33.9	122673	9	US-10-765-790-33	Sequence 33, Appl
c	23	121.6	33.9	571	4	US-09-925-065A-684936	Sequence 684936,
c	24	121.6	33.9	571	4	US-09-925-065A-684938	Sequence 684938,
	25	121.6	33.9	591	5	US-10-027-632-267232	Sequence 267232,
	26	121.6	33.9	591	6	US-10-027-632-267232	Sequence 267232,
c	27	121.4	33.8	29921	8	US-10-719-993-6848	Sequence 6848, Ap
	28	121.4	33.8	106236	8	US-10-741-600-17759	Sequence 17759, A
c	29	121.4	33.8	246144	6	US-10-085-117-226	Sequence 226, App
c	30	121.2	33.8	289	5	US-10-115-278-4	Sequence 4, Appli
c	31	121.2	33.8	289	8	US-10-762-966-4	Sequence 4, Appli
c	32	121.2	33.8	291	6	US-10-229-058B-15	Sequence 15, Appl
c	33	121.2	33.8	571	4	US-09-925-065A-684937	Sequence 684937,
	34	121.2	33.8	3202	4	US-09-925-065A-709882	Sequence 709882,
	35	121.2	33.8	3202	4	US-09-925-065A-709883	Sequence 709883,
c	36	120.8	33.6	498	4	US-09-925-065A-704720	Sequence 704720,
c	37	120.8	33.6	754	5	US-10-027-632-26761	Sequence 26761, A
c	38	120.8	33.6	754	6	US-10-027-632-26761	Sequence 26761, A
c	39	120.8	33.6	779	3	US-09-764-855-191	Sequence 191, App
c	40	120.8	33.6	779	5	US-10-072-349-191	Sequence 191, App
c	41	120.8	33.6	67253	9	US-10-737-082-88	Sequence 88, Appl
c	42	120.8	33.6	67253	9	US-10-765-790-88	Sequence 88, Appl
	43	120.4	33.5	580	4	US-09-925-065A-563601	Sequence 563601,
	44	120.4	33.5	136726	6	US-10-085-117-244	Sequence 244, App
c	45	120.2	33.5	539	4	US-09-925-065A-933530	Sequence 933530,

ALIGNMENTS

RESULT 1
US-10-009-579-5
; Sequence 5, Application US/10009579
; Publication No. US20020156041A1
; GENERAL INFORMATION:
; APPLICANT: Leij de, Lou F.M.H.
; APPLICANT: Ruiters, Marcel H.J.
; APPLICANT: McLaughlin, Pamela M.J.
; APPLICANT: Harmsen, Martin C.
; APPLICANT: Molen v.d., Henk
; APPLICANT: Terpstra, Peter
; APPLICANT: Dokter, Willem H.A.
; TITLE OF INVENTION: Non-squamous epithelium-specific transcription
; FILE REFERENCE: P52075US00
; CURRENT APPLICATION NUMBER: US/10/009,579
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: EP 00200728.4
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/NL01/00166
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(4282)
; OTHER INFORMATION: /note="EGP-2 promotor sequence from -3967 to +315"
US-10-009-579-5

Query Match 100.0%; Score 359; DB 5; Length 4282;
Best Local Similarity 100.0%; Pred. No. 1.5e-75;
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3188 CCCGGCTAATTTTGTATCTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 3247

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3248 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 3307

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3308 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 3367

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3368 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 3427

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTTTTCTTTCTAATTTCAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3428 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTTTTCTTTCTAATTTCAA 3487

Qy        301 AGGAGTATAATTAAATTGCCAGGTAAAAGCTCAAAGGTCTTTTTTATAGTGTTCTGGAA 359
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3488 AGGAGTATAATTAAATTGCCAGGTAAAAGCTCAAAGGTCTTTTTTATAGTGTTCTGGAA 3546


RESULT 2
US-09-925-065A-72033/c
; Sequence 72033, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72033
;   LENGTH: 2041
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-72033

  Query Match             34.5%;  Score 124;  DB 4;  Length 2041;
  Best Local Similarity   64.8%;  Pred. No. 2.1e-19;
  Matches  184;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||     |||||||||| ||||| ||| | ||||| ||||||||||||||
Db       1279 CCCGGCTAATTTTTGTATTTTTAGTAGTGACGGGGTTTCACCATGCTGGTCAGGCTGGTC 1220

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | ||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db       1219 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 1160

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| ||||||||| | ||||    || ||||||  || ||    |  ||    |
Db       1159 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 1100

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
                |    ||||  || |  | || ||   | |      |      |     |||| ||
Db       1099 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 1040

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
                ||| | |||  |  |    | |  | | | | | | | ||
Db       1039 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 996


RESULT 3
US-09-925-065A-72034/c
; Sequence 72034, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72034
;   LENGTH: 2041
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-72034

  Query Match             34.5%;  Score 124;  DB 4;  Length 2041;
  Best Local Similarity   64.8%;  Pred. No. 2.1e-19;
  Matches  184;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||     |||||||||| ||||| ||| | ||||| ||||||||||||||
Db       1279 CCCGGCTAATTTTTGTATTTTTAGTAGTGACGGGGTTTCACCATGCTGGTCAGGCTGGTC 1220

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | ||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db       1219 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 1160

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| ||||||||| | ||||    || ||||||  || ||    |  ||    |
Db       1159 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 1100

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
                |    ||||  || |  | || ||   | |      |      |     |||| ||
Db       1099 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 1040

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
                ||| | |||  |  |    | |  | | | | | | | ||
Db       1039 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 996


RESULT 4
US-09-925-065A-72035/c
; Sequence 72035, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 72035
;   LENGTH: 2041
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-72035

  Query Match             34.5%;  Score 124;  DB 4;  Length 2041;
  Best Local Similarity   64.8%;  Pred. No. 2.1e-19;
  Matches  184;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||     ||||||||| ||||| ||| | ||||| ||||||||||||||
Db       1279 CCCGGCTAATTTTTGTATTTTTAGTAGTGACGGGGTTTCACCATGCTGGTCAGGCTGGTC 1220

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | ||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db       1219 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 1160

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| ||||||||| | ||||     || ||||||  || ||     |  ||     |
Db       1159 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 1100

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
                |    ||||  || |  | || ||   | |     |     |    |||| ||
Db       1099 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 1040

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
                ||| | |||  |  |    | |   | | | | | | | ||
Db       1039 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 996


RESULT 5
US-10-357-930-13830
; Sequence 13830, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel. Robert
;  APPLICANT:  Endege, Wilson
;  APPLICANT:  Monahan, John
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;  TITLE OF INVENTION:  HUMAN PROSTATE CANCER
;  FILE REFERENCE: MRI-007BCN
;  CURRENT APPLICATION NUMBER: US/10/357,930
;  CURRENT FILING DATE:  2003-02-04
;  PRIOR APPLICATION NUMBER: 09/785,276
;  PRIOR FILING DATE: 2003-02-16
;  PRIOR APPLICATION NUMBER: 60/183,319
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: 60/189,862
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: 60/207,454
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/211,314
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/219,007
;  PRIOR FILING DATE: 2000-07-18
;  PRIOR APPLICATION NUMBER: 60/255,281
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 62232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13830
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-357-930-13830

  Query Match             34.5%;  Score 123.8;  DB 8;  Length 438;
  Best Local Similarity   68.8%;  Pred. No. 1.2e-19;
  Matches  170;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              ||||||||||||     |||||||||||| || ||| | ||||||||| ||||||| |||

Db        109 CCGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGATCT 168

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              | |||| |  |||||| ||| ||| ||| ||| |||||||||||||| ||||||||||
Db        169 CAAACTCCTGACCTCATGTGGTCCACCCACCTTGGCCTCCCAAAGTGCTGGGATTACAGG 228

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAATG 181
              ||||||||||||||||||||||    | ||  ||||   | |    ||  |  ||   ||
Db        229 CGTGAGCCACCGCGCTCAGCCTACAAGCAATTTTTTAAAAACTGACACCTATTACTGATA 288

Qy        182 CTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTCC 241
                  |        |||||    |||| |   |  ||| ||| ||  | | |||||  ||
Db        289 AAATTCTCTGTTTAAAAATTCCTATTTACTTTTTTTAAAAAAAAGTATTAATTTTTTTTT 348

Qy        242 CAAGAGA 248
                || | |
Db        349 AAATACA 355


RESULT 6
US-10-027-632-117277/c
; Sequence 117277, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 117277
;   LENGTH: 1136
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-117277

  Query Match             34.4%;  Score 123.6;  DB 5;  Length 1136;
  Best Local Similarity   64.4%;  Pred. No. 2e-19;
  Matches  183;  Conservative    1;  Mismatches  100;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||     ||||||||| ||:|| ||| | ||||| ||||||||||||||
Db        367 CCCGGCTAATTTTTGTATTTTTAGTAGTGAYGGGGTTTCACCATGCTGGTCAGGCTGGTC 308

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | ||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db        307 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 248

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| ||||||||| | ||||     || ||||||  || ||     |  ||     |
Db        247 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 188

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
                |    ||||  || |  | || ||   | |     |     |    |||| ||
Db        187 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 128

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
```

```
                 ||| | |||  |  |    | |   | | | | | | | ||
Db        127 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 84


RESULT 7
US-10-027-632-117278/c
; Sequence 117278, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 117278
;   LENGTH: 1136
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-117278

  Query Match             34.4%;  Score 123.6;  DB 5;  Length 1136;
  Best Local Similarity   64.4%;  Pred. No. 2e-19;
  Matches  183;  Conservative    1;  Mismatches  100;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||     |||||||| ||:|| ||| | ||||| ||||||||||||||
Db        367 CCCGGCTAATTTTTGTATTTTTAGTAGTGAYGGGGTTTCACCATGCTGGTCAGGCTGGTC 308

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | |||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db        307 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 248

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| ||||||||| | ||||    || ||||||  || ||    |  ||    |
Db        247 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 188

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
              |    ||||  || |  | || ||   | |     |     |     |||| ||
Db        187 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 128

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
                ||| | |||  |  |    | |   | | | | | | | ||
Db        127 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 84


RESULT 8
US-10-027-632-117277/c
; Sequence 117277, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 117277
;   LENGTH: 1136
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-117277

  Query Match             34.4%;  Score 123.6;  DB 6;  Length 1136;
  Best Local Similarity   64.4%;  Pred. No. 2e-19;
  Matches  183;  Conservative    1;  Mismatches  100;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||     |||||||| ||:|| ||| | ||||| ||||||||||||||
Db        367 CCCGGCTAATTTTTGTATTTTTAGTAGTGAYGGGGTTTCACCATGCTGGTCAGGCTGGTC 308

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | |||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db        307 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 248

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| ||||||||| | ||||    || ||||||  || ||    |  ||    |
Db        247 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 188

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
              |    ||||  || |  | || ||   | |     |     |     |||| ||
Db        187 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 128

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
                ||| | |||  |  |    | |   | | | | | | | ||
Db        127 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 84


RESULT 9
US-10-027-632-117278/c
; Sequence 117278, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 117278
;   LENGTH: 1136
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-117278

  Query Match             34.4%;  Score 123.6;  DB 6;  Length 1136;
  Best Local Similarity   64.4%;  Pred. No. 2e-19;
  Matches  183;  Conservative    1;  Mismatches  100;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     ||||||||| ||:|| ||| | ||||| ||||||||||||||
Db        367 CCCGGCTAATTTTTGTATTTTTAGTAGTGAYGGGGTTTCACCATGCTGGTCAGGCTGGTC 308

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | ||||||||||||||| ||| ||||||||||||||||||||||||||||
Db        307 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 248

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| ||||||||| | ||||    || ||||||  || ||    |  ||      |
Db        247 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 188

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
                |    ||||  || |  | || ||   | |      |      |     |||| ||
Db        187 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 128

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
              ||| | |||  |  |    | |  | | | | | | | ||
Db        127 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 84


RESULT 10
US-10-417-375-44/c
; Sequence 44, Application US/10417375
; Publication No. US20040219528A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
;  FILE REFERENCE: 529452001600
;  CURRENT APPLICATION NUMBER: US/10/417,375
;  CURRENT FILING DATE:  2003-04-15
;  NUMBER OF SEQ ID NOS: 176
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
;   LENGTH: 53623
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-417-375-44

  Query Match             34.3%;  Score 123;  DB 8;  Length 53623;
  Best Local Similarity   75.4%;  Pred. No. 1.4e-18;
  Matches  153;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              || ||||||||||     |||||||| |||| | ||| | ||||||||| ||||||||||
Db       6062 CCTGGCTAATTTTTGTATTTTTAGTACAGACAGGGTTTCACCATGTTGGCCAGGCTGGTC 6003

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| |  ||||||||||||| |||||||||||||||||||||||| |||||||||
Db       6002 TTGAACTCCTGACCTCAGGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 5943

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||||||||||| |||||     | || |||| |      |||     | | ||
Db       5942 GCGTGAGCCACCGCGCCCAGCCCTCAGAAACATTTTATGTGTATTTCTGCCTCCAAAATA 5883

Qy        181 GCTTATGAAAACGAAAAAAGAAT 203
              || | |||  |  ||||||||
Db       5882 ATTTCTAAAATTGCTAAAAGAAT 5860


RESULT 11
US-10-235-192A-35
; Sequence 35, Application US/10235192A
; Publication No. US20040043389A1
; GENERAL INFORMATION:
;  APPLICANT: McCarthy, Jeanette
;  TITLE OF INVENTION: Methods and Compositions for Identifying
;  TITLE OF INVENTION:  Risk Factors for Abnormal Lipid Levels and the Diseases
;  TITLE OF INVENTION:  and Disorders Associated Therewith
;  FILE REFERENCE: MMI-011
;  CURRENT APPLICATION NUMBER: US/10/235,192A
;  CURRENT FILING DATE:  2002-09-04
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 35
;   LENGTH: 160361
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 151844-151943, 154924-155023
;   OTHER INFORMATION: N = any nucleotide
US-10-235-192A-35

  Query Match             34.1%;  Score 122.4;  DB 7;  Length 160361;
  Best Local Similarity   80.0%;  Pred. No. 3e-18;
  Matches  144;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||||| ||||||||||  || ||| ||||||||||||||||||||||
Db     157896 CCCGGCTAATTTTGTATTTTTTAGTAGAGGTGGGGTTTCTCCATGTTGGTCAGGCTGGTC 157955

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              || |||| |  |||||||||||||||||||||| |||||||||||||| | |||||||||
Db     157956 TCAAACTCCCGACCTCAGGTGATCCGCCCGCCTTGGCCTCCCAAAGTGTTGGGATTACAG 158015

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||| ||||||||||||| | ||   |  | |||  ||||  | | |     | ||   |||
Db     158016 GCGCGAGCCACCGCGCCCGGCTGAGCCAGACCACTTCTGTCTTTTAAGGCTCCTGCTAAT 158075


RESULT 12
US-09-925-065A-591033
; Sequence 591033, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 591033
;   LENGTH: 464
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-591033

  Query Match             34.0%;  Score 122;  DB 4;  Length 464;
  Best Local Similarity   68.6%;  Pred. No. 3.3e-19;
  Matches  181;  Conservative    1;  Mismatches   81;  Indels    1;  Gaps    1;
```

```
Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              |||||||| ||| ||| |||||||| ||| | ||| ||||||||||||||||||||||
Db          1 CCGGCTAACTTTTTAT-TTTTAGTAGCGACAGGGTTTCTCCATGTTGGTCAGGCTGGTCT 59

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
               ||||| | ||||||||||||||||||||||||||||||||||||||||| |||||| |||
Db         60 TGAACTCCCAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTATAGG 119

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAATG 181
              |||||||||| |||| || ||| |||  |  ||| ||   | |   |   || | || |
Db        120 CGTGAGCCACTGCGCCTAGGCTGAGAAAGCTATTTTTTTTTTTTAAGAAAACTGGGAAAG 179

Qy        182 CTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTCC 241
                ||   |   | | |  | || |||| | |   | | |     |    ||  | |||
Db        180 TACATATTATGAAGAGAGTATTTCTTAACATTTTATTTGTGAAGCTTATCTATGTCTGAA 239

Qy        242 CAAGAGAGCCAAGATTTCTTCTTT 265
                : | |||| |  |  ||  |||
Db        240 TCMAACAGCCCACTTGACTCATTT 263


RESULT 13
US-09-925-065A-724096
; Sequence 724096, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 724096
;   LENGTH: 1031
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-724096

  Query Match             34.0%;  Score 122;  DB 4;  Length 1031;
  Best Local Similarity   68.0%;  Pred. No. 4.7e-19;
  Matches  170;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||||| |||||||||||| |  ||| ||||||||||||||||||||||
Db        338 CCCGGCTAATTTTGTATTTTTTAGTAGAGATGCAGTTTCTCCATGTTGGTCAGGCTGGTC 397

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              |||||||||  |||||||||| ||| | ||||||||||||||||||||| ||||||||||
Db        398 TCGAACTTCCGACCTCAGGTAATCTGTCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 457

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||| |||  || | ||||    |      ||   || |  |       |    ||
Db        458 GCGTGAGCAACCATGCCCCGCCTTACTAATAACTTCTATAGAAGTGTTCCATCCTTCAAA 517

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
                || ||      |||| | ||| |   || |   || ||  ||     |||  |   ||
Db        518 ATTTTTGCTGGTTAAAATAAAATAAGACAGTGCTGTTCTATTGAGCATTTCAAATGTATC 577

Qy        241 CCAAGAGAGC 250
              |  ||   ||
Db        578 CACAGTATGC 587


RESULT 14
US-10-741-600-17566/c
; Sequence 17566, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17566
;   LENGTH: 321019
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(321019)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17566

  Query Match             34.0%;  Score 122;  DB 8;  Length 321019;
  Best Local Similarity   76.8%;  Pred. No. 5e-18;
  Matches  149;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              | || |||||||     |||||||||||| || ||| |||||||||||||||||||| |
Db     167635 CTGGTTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTTT 167576

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              |||||| |  |||||||||||||| ||||||| ||||| |||||||| ||||||||||
Db     167575 CGAACTCCTGACCTCAGGTGATCCGGCCGCCTCAGCCTCTCAAAGTGCTGGGATTACAGG 167516

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAATG 181
              |||||||||||||| ||||| | |||    |||||| ||||  |    ||     |
Db     167515 CGTGAGCCACCGCGTCCAGCCAGAGAAAGAATTTTCTAACATTCTAGGTTGTCTACAGAC 167456

Qy        182 CTTATGAAAACGAA 195
              ||||||| | ||
Db     167455 AATATGAAAGCAAA 167442


RESULT 15
US-10-388-838-48/c
; Sequence 48, Application US/10388838
; Publication No. US20040180344A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
;  FILE REFERENCE: 529452001600
;  CURRENT APPLICATION NUMBER: US/10/388,838
;  CURRENT FILING DATE:  2003-03-14
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
;   LENGTH: 329019
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(329019)
;   OTHER INFORMATION: n = A,T,C or G
US-10-388-838-48

  Query Match             34.0%;  Score 122;  DB 8;  Length 329019;
  Best Local Similarity   76.8%;  Pred. No. 5.1e-18;
  Matches  149;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;
```

```
Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              | || |||||||     ||||||||||||| || ||| ||||||||||||||||||||| |
Db     171636 CTGGTTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTTT 171577

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              |||||| |  |||||||||||||| ||||||| ||||| ||||||||| ||||||||||
Db     171576 CGAACTCCTGACCTCAGGTGATCCGGCCGCCTCAGCCTCTCAAAGTGCTGGGATTACAGG 171517

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAATG 181
              |||||||||||||||  ||||| | |||    |||||| ||||  |    ||     |
Db     171516 CGTGAGCCACCGCGTCCAGCCAGAGAAAGAATTTTCTAACATTCTAGGTTGTCTACAGAC 171457

Qy        182 CTTATGAAAACGAA 195
              ||||||| | ||
Db     171456 AATATGAAAGCAAA 171443
```

Search completed: May 6, 2006, 12:21:06
Job time : 578 secs

GenCore version 5.1.8

OM nucleic - nucleic search, using sw model

Run on: May 6, 2006, 12:11:37 ; Search time 423 Seconds
(without alignments)
3456.361 Million cell updates/sec

Title: US-10-009-579A-5_COPY_3188_3546
Perfect score: 359
Sequence: 1 cccggctaatttttgtatctt.....ttttttatagtgttctggaa 359

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9306428 seqs, 2036268586 residues

Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq1:*
2: /SIDS5/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq:*
5: /SIDS5/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
6: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq:*
7: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq1:*
8: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
9: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq:*
10: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
11: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
12: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq3:*
13: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq4:*
14: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq:*
15: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
16: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
17: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
18: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq5:*
19: /SIDS5/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	124	34.5	2041	7	US-09-925-065A-72033	Sequence 72033, A
c	2	124	34.5	2041	7	US-09-925-065A-72034	Sequence 72034, A
c	3	124	34.5	2041	7	US-09-925-065A-72035	Sequence 72035, A
c	4	124	34.5	2041	11	US-10-301-480-173272	Sequence 173272,
c	5	124	34.5	2041	11	US-10-301-480-173273	Sequence 173273,
c	6	124	34.5	2041	11	US-10-301-480-173274	Sequence 173274,
c	7	124	34.5	2041	12	US-10-301-480-786681	Sequence 786681,
c	8	124	34.5	2041	12	US-10-301-480-786682	Sequence 786682,
c	9	124	34.5	2041	12	US-10-301-480-786683	Sequence 786683,
	10	122.2	34.0	63984	17	US-11-121-086-26	Sequence 26, Appl
	11	122	34.0	464	7	US-09-925-065A-591033	Sequence 591033,
	12	122	34.0	1031	7	US-09-925-065A-724096	Sequence 724096,
c	13	122	34.0	321019	10	US-10-995-561-13204	Sequence 13204, A
	14	121.8	33.9	424	7	US-09-925-065A-47240	Sequence 47240, A
	15	121.8	33.9	424	11	US-10-301-480-148478	Sequence 148478,
	16	121.8	33.9	424	12	US-10-301-480-761887	Sequence 761887,
c	17	121.6	33.9	571	7	US-09-925-065A-684936	Sequence 684936,
c	18	121.6	33.9	571	7	US-09-925-065A-684938	Sequence 684938,
	19	121.6	33.9	187745	17	US-11-121-086-83	Sequence 83, Appl
c	20	121.4	33.8	830	12	US-10-301-480-550662	Sequence 550662,
c	21	121.4	33.8	830	12	US-10-301-480-1164071	Sequence 1164071,
c	22	121.2	33.8	571	7	US-09-925-065A-684937	Sequence 684937,
	23	121.2	33.8	3202	7	US-09-925-065A-709882	Sequence 709882,
	24	121.2	33.8	3202	7	US-09-925-065A-709883	Sequence 709883,
c	25	120.8	33.6	498	7	US-09-925-065A-704720	Sequence 704720,
	26	120.4	33.5	580	7	US-09-925-065A-563601	Sequence 563601,
c	27	120.2	33.5	539	7	US-09-925-065A-933530	Sequence 933530,
c	28	120.2	33.5	1257	11	US-10-301-480-38290	Sequence 38290, A
c	29	120.2	33.5	1257	12	US-10-301-480-651699	Sequence 651699,
	30	120.2	33.5	23983	10	US-10-995-561-13491	Sequence 13491, A
c	31	120	33.4	622	7	US-09-925-065A-764373	Sequence 764373,
c	32	120	33.4	1353	7	US-09-925-065A-694967	Sequence 694967,
c	33	120	33.4	1353	7	US-09-925-065A-694968	Sequence 694968,
c	34	120	33.4	1353	7	US-09-925-065A-694969	Sequence 694969,
c	35	119.8	33.4	16082	10	US-10-995-561-13485	Sequence 13485, A
c	36	119.8	33.4	23894	10	US-10-995-561-13320	Sequence 13320, A
c	37	119.6	33.3	978	12	US-10-301-480-581571	Sequence 581571,
c	38	119.6	33.3	978	12	US-10-301-480-1194980	Sequence 1194980,
c	39	119.4	33.3	580	7	US-09-925-065A-104751	Sequence 104751,
c	40	119.4	33.3	583	11	US-10-301-480-204893	Sequence 204893,
c	41	119.4	33.3	583	12	US-10-301-480-818302	Sequence 818302,
c	42	119.4	33.3	1304	7	US-09-925-065A-706696	Sequence 706696,
c	43	119.4	33.3	1304	7	US-09-925-065A-706697	Sequence 706697,
	44	119.4	33.3	167116	17	US-11-121-086-44	Sequence 44, Appl
c	45	119.2	33.2	127943	11	US-10-330-773-628	Sequence 628, App

ALIGNMENTS

RESULT 1
US-09-925-065A-72033/c
; Sequence 72033, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72033
; LENGTH: 2041
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-72033

Query Match 34.5%; Score 124; DB 7; Length 2041;
Best Local Similarity 64.8%; Pred. No. 5.7e+03;
Matches 184; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

Qy 1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
Db 1279 CCCGGCTAATTTTTGTATTTTTAGTAGTGACGGGGTTTCACCATGCTGGTCAGGCTGGTC 1220

```
Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | |||||  | |||||||||||||||  ||| ||||||||||||||||||||||||||||
Db       1219 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 1160

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||  |||||||||  | ||||    || ||||||  || ||    |  ||    |
Db     ,  1159 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 1100

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
              |    ||||  || |  | || ||   | |     |     |     |||| ||
Db       1099 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 1040

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
                ||| | |||  |  |    | |   | | | | | | ||
Db       1039 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 996


RESULT 2
US-09-925-065A-72034/c
; Sequence 72034, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72034
;   LENGTH: 2041
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-72034

  Query Match             34.5%;  Score 124;  DB 7;  Length 2041;
  Best Local Similarity   64.8%;  Pred. No. 5.7e+03;
  Matches  184;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||     ||||||||| ||||| ||| | ||||| ||||||||||||||
Db       1279 CCCGGCTAATTTTTGTATTTTTAGTAGTGACGGGGTTTCACCATGCTGGTCAGGCTGGTC 1220

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | |||||  | |||||||||||||||  ||| ||||||||||||||||||||||||||||
Db       1219 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 1160

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||  |||||||||  | ||||    || ||||||  || ||    |  ||    |
Db       1159 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 1100

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
              |    ||||  || |  | || ||   | |     |     |     |||| ||
Db       1099 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 1040

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
                ||| | |||  |  |    | |   | | | | | | ||
Db       1039 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 996


RESULT 3
US-09-925-065A-72035/c
; Sequence 72035, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72035
;   LENGTH: 2041
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-72035

  Query Match             34.5%;  Score 124;  DB 7;  Length 2041;
  Best Local Similarity   64.8%;  Pred. No. 5.7e+03;
  Matches  184;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||     ||||||||| ||||| ||| | ||||| ||||||||||||||
Db       1279 CCCGGCTAATTTTTGTATTTTTAGTAGTGACGGGGTTTCACCATGCTGGTCAGGCTGGTC 1220

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | |||||  | |||||||||||||||  ||| ||||||||||||||||||||||||||||
Db       1219 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 1160

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              ||||||  |||||||||  | ||||    || ||||||  || ||    |  ||    |
Db       1159 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 1100

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
              |    ||||  || |  | || ||   | |     |     |     |||| ||
Db       1099 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 1040

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
                ||| | |||  |  |    | |   | | | | | | ||
Db       1039 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 996


RESULT 4
US-10-301-480-173272/c
; Sequence 173272, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 173272
;   LENGTH: 2041
;   TYPE: DNA
;   ORGANISM: Homo sapien
```

US-10-301-480-173272

Query Match 34.5%; Score 124; DB 11; Length 2041;
Best Local Similarity 64.8%; Pred. No. 5.7e+03;
Matches 184; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     ||||||||| ||||| ||| | ||||| ||||||||||||||
Db       1279 CCCGGCTAATTTTTGTATTTTTAGTAGTGACGGGGTTTCACCATGCTGGTCAGGCTGGTC 1220

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | ||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db       1219 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 1160

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| ||||||||| | ||||    || ||||||  || ||    |  ||    |
Db       1159 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 1100

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
                |    ||||  || |  | || ||    | |      |      |    |||| ||
Db       1099 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 1040

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
              ||| | |||  |  |    | |  | | | | | | ||
Db       1039 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 996
```

RESULT 5
US-10-301-480-173273/c
; Sequence 173273, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 173273
; LENGTH: 2041
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-173273

Query Match 34.5%; Score 124; DB 11; Length 2041;
Best Local Similarity 64.8%; Pred. No. 5.7e+03;
Matches 184; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     ||||||||| ||||| ||| | ||||| ||||||||||||||
Db       1279 CCCGGCTAATTTTTGTATTTTTAGTAGTGACGGGGTTTCACCATGCTGGTCAGGCTGGTC 1220

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | ||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db       1219 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 1160

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| ||||||||| | ||||    || ||||||  || ||    |  ||    |
Db       1159 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 1100

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
                |    ||||  || |  | || ||    | |      |      |    |||| ||
Db       1099 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 1040

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
              ||| | |||  |  |    | |  | | | | | | ||
Db       1039 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 996
```

RESULT 6
US-10-301-480-173274/c
; Sequence 173274, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 173274
; LENGTH: 2041
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-173274

Query Match 34.5%; Score 124; DB 11; Length 2041;
Best Local Similarity 64.8%; Pred. No. 5.7e+03;
Matches 184; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     ||||||||| ||||| ||| | ||||| ||||||||||||||
Db       1279 CCCGGCTAATTTTTGTATTTTTAGTAGTGACGGGGTTTCACCATGCTGGTCAGGCTGGTC 1220

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | ||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db       1219 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 1160

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| ||||||||| | ||||    || ||||||  || ||    |  ||    |
Db       1159 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 1100

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
                |    ||||  || |  | || ||    | |      |      |    |||| ||
Db       1099 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 1040

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
              ||| | |||  |  |    | |  | | | | | | ||
Db       1039 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 996
```

RESULT 7
US-10-301-480-786681/c
; Sequence 786681, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 786681
; LENGTH: 2041
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-786681

```
  Query Match             34.5%;  Score 124;  DB 12;  Length 2041;
  Best Local Similarity   64.8%;  Pred. No. 5.7e+03;
  Matches  184;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     ||||||||| ||||| ||| | ||||| ||||||||||||||
Db       1279 CCCGGCTAATTTTTGTATTTTTAGTAGTGACGGGGTTTCACCATGCTGGTCAGGCTGGTC 1220

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | ||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db       1219 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 1160

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| |||||||| | ||||    || ||||||  || ||    |  ||    |
Db       1159 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 1100

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
                |    ||||  || |  | || ||   | |      |      |     |||| ||
Db       1099 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 1040

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
              ||| | |||  |  |    | |   | | | | | | | ||
Db       1039 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 996


RESULT 8
US-10-301-480-786682/c
; Sequence 786682, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 786682
;   LENGTH: 2041
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-786682

  Query Match             34.5%;  Score 124;  DB 12;  Length 2041;
  Best Local Similarity   64.8%;  Pred. No. 5.7e+03;
  Matches  184;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     ||||||||| ||||| ||| | ||||| ||||||||||||||
Db       1279 CCCGGCTAATTTTTGTATTTTTAGTAGTGACGGGGTTTCACCATGCTGGTCAGGCTGGTC 1220

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | ||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db       1219 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 1160

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| |||||||| | ||||    || ||||||  || ||    |  ||    |
Db       1159 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 1100

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
                |    ||||  || |  | || ||   | |      |      |     |||| ||
Db       1099 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 1040

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
              ||| | |||  |  |    | |   | | | | | | | ||
Db       1039 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 996


RESULT 9
US-10-301-480-786683/c
; Sequence 786683, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 786683
;   LENGTH: 2041
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-786683

  Query Match             34.5%;  Score 124;  DB 12;  Length 2041;
  Best Local Similarity   64.8%;  Pred. No. 5.7e+03;
  Matches  184;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||||||||||||||     ||||||||| ||||| ||| | ||||| ||||||||||||||
Db       1279 CCCGGCTAATTTTTGTATTTTTAGTAGTGACGGGGTTTCACCATGCTGGTCAGGCTGGTC 1220

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | ||||| | ||||||||||||||| ||| ||||||||||||||||||||||||||||||
Db       1219 TTGAACTCCCAACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGCTAGGATTACAG 1160

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||| |||||||| | ||||    || ||||||  || ||    |  ||    |
Db       1159 GCGTGACCCACCGCGCCCGGCCTCACTACCCCTTTTTCTATATTACAATATGAACAATTA 1100

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
                |    ||||  || |  | || ||   | |      |      |     |||| ||
Db       1099 ATTATGAAAAAATAATATGGCATGATGGGGGGAGCAGAATGGTAGCTCATCATGTTGGAA 1040

Qy        241 CCAAGAGAGCCAAGATTTCTTCTTTCCTCTTCTTTCTTTTTTTT 284
              ||| | |||  |  |    | |   | | | | | | | ||
Db       1039 AGTAGAAAACCAGCAGATAAAATATAGACATGTCTATCTATGTT 996


RESULT 10
US-11-121-086-26
; Sequence 26, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26
;   LENGTH: 63984
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-26

  Query Match             34.0%;  Score 122.2;  DB 17;  Length 63984;
  Best Local Similarity   90.9%;  Pred. No. 6.3e+02;
  Matches  130;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;
```

```
Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              || ||| |||||||||| |||||||||||| || ||| ||||||||||||||||||||
Db      20757 CCTGGCAAATTTTGTATTTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTC 20816

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              ||||||| | |||||||||||||||||| |||||||||||||||||||| |||||||||
Db      20817 TCGAACTCCCAACCTCAGGTGATCCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 20876

Qy        121 GCGTGAGCCACCGCGCTCAGCCT 143
              | ||||||||||||||| | ||||
Db      20877 GGGTGAGCCACCGCGCCCGGCCT 20899


RESULT 11
US-09-925-065A-591033
; Sequence 591033, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 591033
;   LENGTH: 464
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-591033

  Query Match             34.0%;  Score 122;  DB 7;  Length 464;
  Best Local Similarity   68.6%;  Pred. No. 1.7e+04;
  Matches  181;  Conservative    1;  Mismatches   81;  Indels    1;  Gaps    1;

Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              |||||||| ||| ||| |||||||||| ||| | ||| ||||||||||||||||||||
Db          1 CCGGCTAACTTTTTAT-TTTTAGTAGCGACAGGGTTTCTCCATGTTGGTCAGGCTGGTCT 59

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              ||||| | ||||||||||||||||||||||||||||||||||||||||||| |||||| |||
Db         60 TGAACTCCCAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTATAGG 119

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAATG 181
              |||||||||| |||| || ||| |||  |  ||| ||  | |  |   || | || |
Db        120 CGTGAGCCACTGCGCCTAGGCTGAGAAAGCTATTTTTTTTTTTTTAAGAAAACTGGGAAAG 179

Qy        182 CTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTCC 241
                ||   |  | | |  | || |||| | |   | | |    |     ||  | |||
Db        180 TACATATTATGAAGAGAGTATTTCTTAACATTTTATTTGTGAAGCTTATCTATGTCTGAA 239

Qy        242 CAAGAGAGCCAAGATTTCTTCTTT 265
              : | |||| |  |  ||  |||
Db        240 TCMAACAGCCCACTTGACTCATTT 263


RESULT 12
US-09-925-065A-724096
; Sequence 724096, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 724096
;   LENGTH: 1031
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-724096

  Query Match             34.0%;  Score 122;  DB 7;  Length 1031;
  Best Local Similarity   68.0%;  Pred. No. 1e+04;
  Matches  170;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              |||||||||||||||||| |||||||||||| |  ||| ||||||||||||||||||||
Db        338 CCCGGCTAATTTTGTATTTTTTAGTAGAGATGCAGTTTCTCCATGTTGGTCAGGCTGGTC 397

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              |||||||||  |||||||||| ||| | ||||||||||||||||||||||| |||||||||
Db        398 TCGAACTTCCGACCTCAGGTAATCTGTCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 457

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAAT 180
              |||||||| |||  || | ||||     |      ||   || | |        |    ||
Db        458 GCGTGAGCAACCATGCCCCGCCTTACTAATAACTTCTATAGAAGTGTTCCATCCTTCAAA 517

Qy        181 GCTTATGAAAACGAAAAAAGAATTATTAAGAGTAATTATAAAGAAACACTCATTTTCTTC 240
              || ||      |||| | ||| |   || |   || ||  ||      ||| |  ||
Db        518 ATTTTTGCTGGTTAAAATAAAATAAGACAGTGCTGTTCTATTGAGCATTTCAAATGTATC 577

Qy        241 CCAAGAGAGC 250
              |  ||   ||
Db        578 CACAGTATGC 587


RESULT 13
US-10-995-561-13204/c
; Sequence 13204, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13204
;   LENGTH: 321019
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(321019)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-995-561-13204

  Query Match             34.0%;  Score 122;  DB 10;  Length 321019;
```

```
  Best Local Similarity   76.8%;  Pred. No. 2.1e+02;
  Matches  149;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          2 CCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTCT 61
              | || |||||||     |||||||||||| || ||| ||||||||||||||||||||||| |
Db     167635 CTGGTTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTTT 167576

Qy         62 CGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAGG 121
              |||||| |  ||||||||||||| ||||||| ||||| ||||||||| ||||||||||
Db     167575 CGAACTCCTGACCTCAGGTGATCCGGCCGCCTCAGCCTCTCAAAGTGCTGGGATTACAGG 167516

Qy        122 CGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTTACATCTTCAAGTGCTAGAAATG 181
              |||||||||||||  ||||| | |||    ||||| |||| |  ||    |
Db     167515 CGTGAGCCACCGCGTCCAGCCAGAGAAAGAATTTTCTAACATTCTAGGTTGTCTACAGAC 167456

Qy        182 CTTATGAAAACGAA 195
              ||||||| | ||
Db     167455 AATATGAAAGCAAA 167442


RESULT 14
US-09-925-065A-47240
; Sequence 47240, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47240
;   LENGTH: 424
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-47240

  Query Match             33.9%;  Score 121.8;  DB 7;  Length 424;
  Best Local Similarity   84.9%;  Pred. No. 1.8e+04;
  Matches  135;  Conservative    1;  Mismatches   23;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||||||| |||||||||||||| ||| |||||||||| ||||||||||
Db         56 CCCAGCTAATTTTGTATTTTTTAGTAGAGACGGGGTTTCTCCATGTTGTTCAGGCTGGTC 115

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | |||||| |  |||||||||||||| || |||| ||||||||||||||||||||||||
Db        116 TTGAACTCCTGACCTCAGGTGATCCACCTGCCTTGGCCTCCCAAAGTGCTAGGATTACAG 175

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTT 159
              || ||||||||||:||   ||||||   |   |||| ||
Db        176 GCATGAGCCACCGYGCCTGGCCTGGATTCTGTTTTTATT 214


RESULT 15
US-10-301-480-148478
; Sequence 148478, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 148478
;   LENGTH: 424
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-148478

  Query Match             33.9%;  Score 121.8;  DB 11;  Length 424;
  Best Local Similarity   84.9%;  Pred. No. 1.8e+04;
  Matches  135;  Conservative    1;  Mismatches   23;  Indels    0;  Gaps    0;

Qy          1 CCCGGCTAATTTTGTATCTTTTAGTAGAGACGGCGTTCCTCCATGTTGGTCAGGCTGGTC 60
              ||| |||||||||||||| |||||||||||||| ||| |||||||||| ||||||||||
Db         56 CCCAGCTAATTTTGTATTTTTTAGTAGAGACGGGGTTTCTCCATGTTGTTCAGGCTGGTC 115

Qy         61 TCGAACTTCAAACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACAG 120
              | |||||| |  |||||||||||||| || |||| ||||||||||||||||||||||||
Db        116 TTGAACTCCTGACCTCAGGTGATCCACCTGCCTTGGCCTCCCAAAGTGCTAGGATTACAG 175

Qy        121 GCGTGAGCCACCGCGCTCAGCCTGGGAACACCTTTTCTT 159
              || ||||||||||:||   ||||||   |   |||| ||
Db        176 GCATGAGCCACCGYGCCTGGCCTGGATTCTGTTTTTATT 214


Search completed: May  6, 2006, 12:28:15
Job time : 425 secs
```